

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: February 26, 2002, 07:55:20 ; Search time 28.78 Seconds
(without alignments)
2810.584 Million cell updates/sec

Title: US-09-605-783A-113
Perfect score: 2861
Sequence: 1 MVQRLWVSRLRHRKAQLL.....AIYFATQVWFKSLAKYSA 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB ID	Description
1	519.5	18.2	599	5	Q9VSV1
2	488.5	17.1	530	4	Q9UMX9
3	347.5	12.1	515	10	O65803
4	347.5	12.1	515	10	Q9FNR6
5	346.5	12.1	501	10	Q9SQK6
6	346.5	12.1	515	10	Q9SLN7
7	346	12.1	612	10	Q9SQK5
8	338	11.8	512	10	Q9SQK5
9	337.5	11.8	523	10	O04077
10	336.5	11.8	501	10	Q9SP63
11	333	11.6	512	10	Q9ZTB9
12	330	11.5	516	10	Q43653
13	329.5	11.5	524	10	Q9XHL6
14	328.5	11.5	594	10	O80605
15	325.5	11.4	604	10	Q9FVL6
16	323.5	11.3	512	10	Q9ZP23
17	323.5	11.3	512	10	O80550
18	321.5	11.2	502	10	Q9SP14
19	317.5	11.1	507	10	Q40583

20	317.5	11.1	530	10	Q9M535
21	316.5	11.1	500	10	Q9FVJ6
22	313	10.9	521	10	O9SXM0
23	311	10.9	533	10	O41152
24	307.5	10.7	488	10	Q9FV32
25	306	10.7	523	10	Q9M422
26	305.5	10.7	501	10	O65929
27	302.5	10.6	510	10	O9SP15
28	301.5	10.5	429	10	O40167
29	299.5	10.5	505	10	Q9SQK4
30	296.5	10.4	510	10	Q9M3R4
31	294.5	10.3	510	10	O9FES9
32	294	10.3	510	10	O40938
33	294	10.3	523	10	Q9M438
34	292.5	10.2	512	10	O9C8X2
35	292	10.2	537	10	O49838
36	292	10.2	538	10	Q9LKH3
37	291.5	10.2	509	10	O9FNH9
38	291	10.2	618	4	Q9ULU3
39	287	10.0	539	10	Q9SM24
40	286	10.0	474	10	O04516
41	284.5	9.9	334	10	O65883
42	281.5	9.8	491	10	Q9C6H8
43	281	9.8	491	10	O9FG00
44	280.5	9.8	492	10	Q9FIX9
45	279.5	9.8	513	10	Q39232

ALIGNMENTS

RESULT	1
Q9VSV1	
ID	Q9VSV1 PRELIMINARY; PRT; 599 AA.
AC	Q9VSV1:
DT	01-MAY-2000 (TREMBLrel. 13, Created)
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE	CG4484 PROTEIN.
GN	CG4484.
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.
OX	NCBI_TaxID=7227;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=BERKELEY;
RX	MEDLINE=20196006; PubMed=10731132;
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Ananadides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA	Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.D.G.,
RA	Abriel J.F., Agbayani A., An H.-J., Andrews-Fannkuch C., Baldwin D.,
RA	Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA	Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA	De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA	Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck C.,
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA	Merkulov G., Milshina N.V., Mobarri J., Morris J., Moshrefi A.,

RC STRAIN-SHIRAZ; TISSUE-BERRY;
RA Davies C., Wolf T., Robinson S.P.;
RT "Three putative sucrose transporters are differentially expressed in
RT grapevine tissues."
RL Plant Sci. 147:93-100(1999).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL: AF021808; AAF08329.1; -.
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 501 AA; 53938 MW; 4D0D4DE2EF2F4BA8 CRC64;

Query Match 12.1%; Score 346.5; DB 10; Length 501;
Best Local Similarity 27.0%; Pred. No. 2e-16; Indels 133; Gaps 22;
Matches 153; Conservative 83; Mismatches 197; Indels 133; Gaps 22;
QY 17 QLLLVNLLTFGLVCLAAAGITYVPPPLLEVGVEEKMVTMLVIGIPVLGVCVPLLGASD 76
Db 29 RLLRVASVAGIQCWALQSLTTPYVQELGIPHAWSIIWLCGLSGLLVQPLVGLHSD 88
QY 77 HWRGRRRRPPIWALSIGILLSLFLIPRAGWLAGLL--CPDPRPLEALLILGVLGLD 134
Db 89 RCRSRRRRPPIVAGATSIYVAVLIIGFSADIGGLGGADRRPRAVATFVVGFWLLDV 148
QY 135 CGQVCFPTLEALLSDFRDPD--CRQAYSVYAFMISLGGCLGYLLPA-----IDWDT 186
Db 149 ANNVQGPCRALLADL-TEKHRRTPVANAYFSLFIANGVNLGPFATGSGWFRFWETS 207
QY 187 ALAPYLGTQEECL----FGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPSLPHCC 242
Db 208 TSS----CNADCANLKSFAFLDIIF--IAITTIISITAA-----QELPLSSSRSTHIS 255
QY 243 -----PCRLARLAFNLGALLPRHLCCRMPTLR-----LFVAELCSWALMTFT 289
Db 256 EWAESTHAQEA-----LWELF-----GTLRYLSSGSIWLLFVYAL-TWIGLLPFL 301
QY 290 LFYDFVGEGLYGVPRAPGPTGTEARRHYDEGVRMGSLGFLQCAISLVSFLMDRLVQRF 349
Db 302 LFEDTWMGREIYGGKP-----NEGQNTYGVGMGALGMLNSVVLGITSVLMEKLCRKW 355
QY 350 CTYRAYVLASVAAPFAAGATCLSHSVAVVATASAAITGTFALQILPTLASLYHREKQV 409
Db 356 GAGFVW-----GLSNILMSLCFLMLTILS---AVVRHMD--- 386
QY 410 FLPKYRGDTGGASSEDLSMTSFLPGPKGAPFNGHVAGGSGLLPPPPALCGASACDVS 469
Db 387 -----FLGHDLP-----PSGVVIA-----ALIVFSILGIP 411
QY 470 VRVVVGEV---TEARVP-GRGICLDLALDSAFLLSOVAPSLFMSIVQL-----SOSV 520
Db 412 LAITYSVYALISTRIESLGLGQGLSGMVLNLAIVQIVIVSLGSGPMDQLFGGNGSPSL 471
QY 521 TAYMYSAGLGVAIVYFATQVVDKS 546
Db 472 AVAAVAAPAFASGLVAILAIPRSADKS 497

RESULT 6
Q9SLN7 PRELIMINARY; PRT; 515 AA.
AC Q9SLN7
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE SUCROSE TRANSPORTER PROTEIN.
GN CSUT.
OS Daucus carota (Garrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;

RN SEQUENCE FROM N.A.
RC TISSUE=ROOT;
RA Wu N., Diao F., Zhang L., Huang M.;
RT "Sucrose transporter protein."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL: AB036758; BAA89458.1; -.
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 515 AA; 54523 MW; 2B548D1D9DFF51AC CRC64;

Query Match 12.1%; Score 346.5; DB 10; Length 515;
Best Local Similarity 27.2%; Pred. No. 2e-16; Indels 75; Gaps 19;
Matches 137; Conservative 76; Mismatches 215; Indels 75; Gaps 19;
QY 17 QLLLVNLLTFGLVCLAAAGITYVPPPLLEVGVEEKMVTMLVIGIPVLGVCVPLLGASD 76
Db 34 KLVLVAIAAGVQFGWALQSLTTPYVQELGIPHKWAIWLCGPISGMLVQPIVGYSD 93
QY 77 HWRGRRRRPPIWALSIGILLSLFLIPRA---GWLAG-LCCPDRPLEALLILGVLGL 132
Db 94 HCSFRRRRPPIASGAGCAISVILIGFAADIGYKAGDDMSKTLKPRAVTFVIGFWIL 153
QY 133 DFCQVCFPTLEALLSDFLR-DPDHCRQAYSVAFMISLGGCLGY-----LLPAI 181
Db 154 DVANNMLOGPCRALLADLCSGDTTRMRKSANAFYFFWAVGNILGYAAGSYNNLYKLPPFS 213
QY 182 DWDTSALAPYLGTQEECL--FGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240
Db 214 --KTHACDLYCANLKSFIISALIIITVVA--LSVVRKQWSP-DDADADEP----- 263
QY 241 CCPCRARLAFNLGALLPRHLCCRMPTLRRLFVAELCSWALMTFTLYTDFVGEGL 300
Db 264 --PSGKIPV--FGLIGALKDL---PRMLLLIVTCLNIAWFFFLFDTDMGREGI 315
QY 301 YQGVPRAPGPTGTEARRHYDEGVRMGSLGFLQCAISLVSFLMDRLVQRFCTRAYVLASVA 360
Db 316 YGGT--AGKG---KLYDQGVRAAGSLGILLNSVVLGTSIAVEYLVRGVGVKI-LWGVV 368
QY 361 AFPVAAGATCL-----SHSV-----AVVTASA-----ALTGTFSLQILPY 397
Db 369 NFILAIGLVMTVVYKVAQHOREHSANGQLLPPSAGVKGAGALSLSILGIPLSITYSIPF 428
QY 398 TLASLYHREKQVFLPKYRGDTGGASSEDLSMTSFLPGPKG-----APPNGHVAGG 450
Db 429 ALASISGSGGAGOGSLGVNLNLAIVVPMIVSVLAGFFDSLFGGGLNLPFVVGAIASAI 488
QY 451 SGLLP----PPPALCGASACDVS 469
Db 489 SGVLAIVLLPKPSKDAASKLSLS 511

RESULT 7
Q9SQK5 PRELIMINARY; PRT; 612 AA.
AC Q9SQK5
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE PUTATIVE SUCROSE TRANSPORTER.
GN VVSUG12.
OS Vitis vinifera (Grape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;
OC Vitis.
OX NCBI_TaxID=29760;
RN SEQUENCE FROM N.A.
RP STRAIN-SHIRAZ; TISSUE-BERRY;
RC

RA Davies C., Wolf T., Robinson S.P.;
RT "three putative sucrose transporters are differentially expressed in
RL grapevine tissues";
PL Plant Sci. 147:93-100(1999).
DR EMBL; AF021809; AAF08330.1; -;
SQ SEQUENCE 612 AA; 65711 MW; 8FC3FBD6AB439078 CRC64;

Query Match 12.1%; Score 346; DB 10; Length 612;
Best Local Similarity 25.4%; Pred. No. 2.7e-16;
Matches 125; Conservative 75; Mismatches 178; Indels 114; Gaps 17;

Qy 18 LLLVNLFTGFEVLAAGITYVPPLLLEVGVEEFMTVMVIGVGLVGVPLVGLSASDH 77
Db 18 LLLVNLFTGFEVLAAGITYVPPLLLEVGVEEFMTVMVIGVGLVGVPLVGLSASDH 77
Db 66 LILSOMTAAGVQFGWALQSLTTPYIQTGLIEHAFSSFIWLCGPIGLVQPCVGIWSDK 125
Qy 78 WRGRYGRRRPFIWALSIGILLISFLI---PRAGWLAGLCCDDPRPLE-----LALLILG 128
Db 78 WRGRYGRRRPFIWALSIGILLISFLI---PRAGWLAGLCCDDPRPLE-----LALLILG 128
Db 126 CSSYGRRRPFIWALSIGILLISFLI---PRAGWLAGLCCDDPRPLE-----LALLILG 185
Qy 129 VGLLDFCGVQCFTPLEALLSDFRDPDCHQAOYSVYAFMISLGGCLGYLLLPALID---WD 184
Db 129 VGLLDFCGVQCFTPLEALLSDFRDPDCHQAOYSVYAFMISLGGCLGYLLLPALID---WD 184
Db 186 FWMLDLANNVQGFARALLADL-SCPDORNSANAIFCSWMAVGNILGFSAGSGHWRW- 243
Qy 185 TSALAPYLGTOEEC-----LFGLLTLFIUT-CVAATLLVAEEAALGPTPEAE-GLSAP 235
Db 185 TSALAPYLGTOEEC-----LFGLLTLFIUT-CVAATLLVAEEAALGPTPEAE-GLSAP 235
Db 244 ----FPFLNKKACCEAGCNKAAFLIAVFLTCLTLVTPAEVPLMAYOPHLSDSAP 299
Qy 236 SL-SPHCCPCARLAFLNLI-----LPRMLMLVVTCLNMIWAFVFFILFDDMAGRE 311
Db 236 SL-SPHCCPCARLAFLNLI-----LPRMLMLVVTCLNMIWAFVFFILFDDMAGRE 311
Db 300 LLDNPNQ-----QIGFDSKSLDMSAVDNATGNPNPSSYEINKNAKHLTPIVQEONSF 353
Qy 254 ----GALLPRHLQCCRPRTLRLRFLVAELCSWMALMTFTLYTDFVGEGLYQGVPRAP 309
Db 254 ----GALLPRHLQCCRPRTLRLRFLVAELCSWMALMTFTLYTDFVGEGLYQGVPRAP 309
Db 354 SDGPGAVLVNLTSLRHLPPAMHSVLLVWALSLSWPFPLFDDMAGREYVHGDPKGD- 412
Qy 310 GTEARRHYDEGRVMSGLFLQCAISLVSFLVMDRLVQRFCTRAYV-LASVAAPVPAAGA 368
Db 310 GTEARRHYDEGRVMSGLFLQCAISLVSFLVMDRLVQRFCTRAYV-LASVAAPVPAAGA 368
Db 413 -ESAVKAYDAGREGAFGLLNSVDLGISSFLIEPQCORMGARLVWAMSNIIVACMAGT 471
Qy 369 TCLS-----HSV-----AVVTASA---ALTGFTFSAQLIPLYTLASLYHREKQ 408
Db 413 -ESAVKAYDAGREGAFGLLNSVDLGISSFLIEPQCORMGARLVWAMSNIIVACMAGT 471
Qy 409 VFLPKYRGDTGG 420
Db 409 VFLPKYRGDTGG 420
Db 525 ----ELTADTGG 532

RESULT 8
Q9S7Z5 PRELIMINARY; PRT; 512 AA.
ID Q9S7Z5
AC Q9S7Z5
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SUCROSE TRANSPORTER SUT2B.
GN SUT2B OR SUT2A.
OS Apium graveolens (Celery).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Apium.
OX NCBI_TaxID=4045;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-MATURE LEAF;
RA Noiraud N., Lemoine R.;
RT "Sucrose transporters in celery";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; AF167416; AAD45391.1; -;
DR EMBL; AF167415; AAD45390.1; -;

DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 512 AA; 54426 MW; 2637535216FF1ED2 CRC64;
Query Match 11.8%; Score 338; DB 10; Length 512;
Best Local Similarity 25.3%; Pred. No. 7.9e-16;
Matches 140; Conservative 87; Mismatches 198; Indels 128; Gaps 21;
Qy 17 QLLVNLFTGFEVLAAGITYVPPLLLEVGVEEFMTVMVIGVGLVGVPLVGLSASD 76
Db 17 QLLVNLFTGFEVLAAGITYVPPLLLEVGVEEFMTVMVIGVGLVGVPLVGLSASD 76
Db 33 KLILVAIAAGVQFGWALQSLTTPYIQTGLIEHAFSSFIWLCGPIGLVQPCVGIWSD 92
Qy 77 WRGRYGRRRPFIWALSIGILLISFLI---GWLAG-LCCPDPRPLEALLILGVL 132
Db 77 WRGRYGRRRPFIWALSIGILLISFLI---GWLAG-LCCPDPRPLEALLILGVL 132
Db 93 RCSSFGRRPFIWALSIGILLISFLI---GWLAG-LCCPDPRPLEALLILGVL 152
Qy 133 DFCQVQCFTPLEALLSDFRDPDCHQAOYSVYAFMISLGGCLGY-----LLPAI 181
Db 133 DFCQVQCFTPLEALLSDFRDPDCHQAOYSVYAFMISLGGCLGY-----LLPAI 181
Db 153 DVANNMLQGPCRALLADLCLNGDTRMRSAFYSFFMAVGNILGYAAGSYNNLYKLPFFS 212
Qy 182 DWDTALAPYLGTOEEC-LFGLTLFIUT-CVAATLLVAEEAALGPTPEAEGLSAPSLSPH 240
Db 182 DWDTALAPYLGTOEEC-LFGLTLFIUT-CVAATLLVAEEAALGPTPEAEGLSAPSLSPH 240
Db 213 --KTHACDLVCANLKSCFIISIVLITVLT--VREKORSDEADEEPPSSGKIP- 267
Qy 241 CCPCARLAFNL-GALLPRHLQCCRPRTLRLRFLVAELCSWMALMTFTLYTDFVGE 299
Db 241 CCPCARLAFNL-GALLPRHLQCCRPRTLRLRFLVAELCSWMALMTFTLYTDFVGE 299
Db 268 ----VFGLFGALKD-----LPRMLMLVVTCLNMIWAFVFFILFDDMAGRE 311
Qy 300 LYQGVPRAEPTARRHYDEGRVMSGLFLQCAISLVSFLVMDRLVQRFCTRAYVLA 359
Db 300 LYQGVPRAEPTARRHYDEGRVMSGLFLQCAISLVSFLVMDRLVQRFCTRAYVLA 359
Db 312 IYGT--AGOG---KLYDQGVMSGLLNSVVLGTSIAVEYLVRGVGGVKI-LWGL 364
Qy 360 AAFVPAAGATCLSHSAVAVTASAALTGFTFSAQLIPLYTLASLYHREKQVFLPKY 419
Db 360 AAFVPAAGATCLSHSAVAVTASAALTGFTFSAQLIPLYTLASLYHREKQVFLPKY 419
Db 365 VNFLLAIGLV-----MTVWVSKVA-----QHQHQH----- 389
Qy 420 GASSEDLSMTSFLPCPKPGAPFNGHVAGSGGLLPPPPALCGASACDVSRVVVVG 479
Db 420 GASSEDLSMTSFLPCPKPGAPFNGHVAGSGGLLPPPPALCGASACDVSRVVVVG 479
Db 390 ----GANGOLLPSA--GVKAGALSFLSILGILS 418
Qy 480 ARV-VP-----GRGICLDLAILDSAFLLSQVAPSLFMGSIVOL--SQSVTAYMV 525
Db 480 ARV-VP-----GRGICLDLAILDSAFLLSQVAPSLFMGSIVOL--SQSVTAYMV 525
Db 419 ITFSIPALASIYSSGSGAGGSLGVNLAIIVVPMIVSVLAGPFDLSLFGGGLPAFVV 478
Qy 526 ---SAAGLGLVAI 535
Db 526 ---SAAGLGLVAI 535
Db 479 GAISAAISGVLA 491

RESULT 9
O04077 PRELIMINARY; PRT; 523 AA.
ID O04077
AC O04077
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SUCROSE TRANSPORT PROTEIN.
GN SUT
OS Vicia faba (Broad bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Vicia.
OX NCBI_TaxID=3906;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-COTYLEDON;
RA Weber H., Borisjuk L., Heim U., Sauer N., Wobus U.;
RL Plant Cell 0:0-0(0).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; Z93774; CAB07811.1; -;

```
DR Mendel: 16626; Vicfa:2554:16626.
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 523 AA; 55229 MW; F8EBE170212D191A CRC64;

Query Match      11.8%; Score 337.5; DB 10; Length 523;
Best Local Similarity 25.8%; Pred. No. 8.8e-16;
Matches 134; Conservative 76; Mismatches 185; Indels 125; Gaps 20;

QY 17 QLLVNLTLTGLVCLAAAGITVVPPLLEVGVEEKFMTWLGIGPVGLGVCPILGASD 76
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 36 KIMVASTAAGVQFWALQSLPIYPVOLLGIHTWAAIYIWLCPGISMLVQPIVGYISD 95
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 77 HWRGRRRRPFIFWALSIGILLSLFLIPRAGWLAGL-----LCPDRPLELALLILGV 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 96 RCTSRFRRRPFIAAGSTAVAI AFLI--GYAADLGHSGFSDLDQKVRPRAIGIFVVG 152
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 130 GLLDFCGQVCTFPLEALLSDIFR-DPDHCRQAYSVYAFMISLGCGCLGYLLPAID----- 182
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 153 WILDVANNMLOGPCRALLDGLDCAGNQRTRNANAFFSFMAVGNVLYGAAGAYS KLYHVF 212
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 183 --WDTSAALPYLGTQOECLFGLLTFLTCVAATLLV-AEEAALGPTPEA---EGLSAP 235
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 213 PPTKACNVKCANLKSCFP--LSIALTLVLTATSAIIVVKETALTPEKTVVTTEDGSGG 270
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 236 SLSPHCCPCRLR--AFRNLGALLPRLHOLCCRPRTLRLRFLVAELCSWMAIMTFTLFT 293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 271 GM-----PCFGSLGAFREL-----KRPWMLLLVTCLNWIAWFFELLFDT 311
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 294 DFVGEGLVQVPRAEPGTEARRH-YDEGVRMGSLGLFLOCAISLVFSLVMDRLVQRF 352
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 312 DMWGKEVY-----GGTVGEGHAYDMGVREGALGLMLNSVVLGATSLGVLDILARGVG-G 363
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 353 AVYLASVAAPVAAGATCLSHSVAVV-----LFTVAELCSWMAIMTFTLFT 387
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 364 VKRLWGIWVFL--AICLGIITLVTKLAHSRQVAPGTGALGPLPPEGIKAGALTFL 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 388 TFSALQI-----LPYTLASLYHREKQVFLPKYRGDTGGASSEDLSL-----MTS 430
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 421 SVLGVPLAITYSIPALASIF-----SSTSGAGGSLGLVNLNLAIVIPQMFVS 468
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 431 FLPGPKPG-----APFNGHVAGGSG-----LLPPPP 458
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 469 VLSGPDALFGGNLPAPFVWGAVAALASGILSIILLSPSP 508
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
Q9SP63 PRELIMINARY; PRT; 501 AA.
AC Q9SP63;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SUCROSE TRANSPORTER.
OS Vitis vinifera (Grape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;
OC Vitis.
OX NCBI_TaxID=29760;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. UGNI BLANC; TISSUE=GRAPE BERRY;
RA Ageorges A., Delrot S., Romieu C.;
RT "Sucrose transporter from grape berry.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL: AF182445; AAD5289.1; -.
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
KW Transmembrane.

SQ SEQUENCE 501 AA; 54075 MW; 4FD19DFAFE539077 CRC64;

Query Match      11.8%; Score 336.5; DB 10; Length 501;
Best Local Similarity 26.9%; Pred. No. 9.8e-16;
Matches 152; Conservative 82; Mismatches 199; Indels 133; Gaps 22;

QY 17 QLLVNLTLTGLVCLAAAGITVVPPLLEVGVEEKFMTWLGIGPVGLGVCPILGASD 76
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 29 LLRVASVAGGIQFWALQSLTPIYVQELGIPHAWSSIIWLCPGLSGLLVQPIVGLHLS 88
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 77 HWRGRRRRPFIFWALSIGILLSLFLIPRAGWLAGL--CPDRPLELALLILGVGLDF 134
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 89 RCNSRFRRRPFIVAGATSIIVAVLIIGTFDGLGDCGADRRPRAVATVVGFWLLDV 148
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 135 CGQVCTFPLEALLSDIFRDPH--CROAYSVYAFMISLGCGCLGYLLPA-----IDMTS 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 149 ANNVTOGPCRALLADL-TEKHRRTRVANAYFSLFIAVGNVLFATGTSYSGWERIFWFTS 207
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 187 ALAPYLTQBECL-----FGLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSPHCC 242
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 208 TSS-----CNADCANLKSAFLDIIF--IAITVISITAA-----QELPLSSSSSRSTHIS 255
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 243 -----PCRARLAPRNLGALLPRLHOLCCRPMTLRR-----LFVAELCSWMAIMTFT 289
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 256 EEMAESTHAQEA-----LWELF-----GTLRYLSGSIWIILFYTAL-TWIGWFFL 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 290 LFYTFYFVGEGLVQVPRAEPGTEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVQRF 349
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 302 LFDTDMWGRIYGGKP-----NEGQNVNTGVRMGALGLMLNSVVLGITSVLMEKLCRKW 355
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 350 GTRAVYLASVAAPVAAGATCLSHSVAVVASAALTGFTFSALQILPYTLASLYHREKQV 409
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 356 GAGEFW-----GLSNTILMSLCFLMLILS---AVVKHMD--- 386
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 410 FLPKYRGDTGGASSEDLSMTSFLPGPKGAPPNGHVAGGSGLLPPPPALCGASACDVS 469
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 387 -----FLGHDLP-----PSGVVIA-----ALIVFSILGIP 411
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 470 VRVVGPE---TEARVVP-GRGICIDLAIDSAFLSQVAPSLPMGSIQV-----SQSV 520
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 412 LAITYSPYALISTRIESLGLGOGLSMGVLNLAIVIPQIVTSLGSGPDQLFGGNGSPSL 471
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 521 TAYMVSAGLGLVAIYATQVWFDS 546
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 472 AVAAVAFASGLVAILAIPRSSADKS 497
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
Q9ZTB9 PRELIMINARY; PRT; 512 AA.
AC Q9ZTB9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SUCROSE TRANSPORTER.
GN SUT1.
OS Apium graveolens (Celery).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Apium.
OX NCBI_TaxID=4045;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RA Noiraud N., Delrot S., Lemoine R.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL: AF063400; AAC99332.1; -.
DR Mendel: 36163; Apigr:2554;36163.
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
```



```
SQ SEQUENCE 524 AA; 55279 MW; B5CE3F880D2C1E6B CRC64;

Query Match 11.5%; Score 329.5; DB 10; Length 524;
Best Local Similarity 25.5%; Pred. No. 3.2e-15;
Matches 133; Conservative 74; Mismatches 187; Indels 127; Gaps 19;

QY 17 QLLVNLITFGLEVCLAAGITYPPPLLELVGEVEEKFTMTVVLGIPVLGVCVPLLGASD 76
DB 37 KIMVVASIAAGVFGWALQSLTTPYVQLLGIHTTAAAYIWLGCPIGSLMLVQVGVYHSD 96
QY 77 HWRGRYGRRRPFTWALSGLTLLSLFIPRAGWLAGL-----LCPDPRPLELALLIGV 129
DB 97 RCTSRGRRPFTAGSIAVAIAVELI---GYAADLGHSGDNLDKKVRPRAIGIFVGVF 153
QY 130 GLIDFCQCVFTPLEALLSDLFR-DPDHCRQAYSVAFMISLGGCLGCLYLLPAID----- 182
DB 154 WILDVANNMLQGPCCRALLDGLCAGNQRKTRNANAFSFFMAVGNVLGYAAGAYSPLYVVF 213
QY 183 --WDTSAALPYLGTQECFLGLTLTFLTCVAA-----TLVAEEAALGPTPEAGLSA 234
DB 214 PFTKTEACNYCANLKSFCFLSTALITVLTATAALIYVKETPLIAERAVV---TAEDGGSN 270
QY 235 PSLSPHCCPCARL--AFRNLGALLPRLHOLCCMRPTLRLRVAELCSWMLMTFTLFY 292
DB 271 GGM-----PCFGOLSGAFKEL-----KRPMMILLVTLNWIAMFPFLFD 311
QY 293 TDFVGBGLVQGVPRAPGPGTEARRH-YDEGVRMGSLGLFLQCAISLVSFLVMDRLVQFCT 351
DB 312 TDMMGKEVY-----GGTVGEHAYDMGVRAAGLGLMLNSVVLGATSLGVDILARGV- 363
QY 352 RAVLVASVAPVAAAGATCLSHSVAVV-----TASAAALG 386
DB 364 GVRLMGIVNELL--AICLGLFVLTKLAQHSQYAPGTGGLODPLPPSGGKAGALTL 420
QY 387 FTSAIQI-----LPTYLASLYHREKQVLPKYRGDTGASSEDLSL-----MT 429
DB 421 FSVLGIPLAITYSIPFALASIF-----SSTSGAGQGLSLGLVNLAIIVPQMFV 468
QY 430 SFLPGPKPG-----APPNHGHVAGGSG-----LLPPPP 458
DB 469 SVLSGPDWALFGGGLNLPFAFVGVGAALASGILSMILLPSP 509

RESULT 14
ID 080605 PRELIMINARY; PRT; 594 AA.
AC 080605;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE PUTATIVE SUCROSE/H+ SYMPORTER (SUCROSE TRANSPORTER).
GN T17M13.3 OR SUC3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RT "Arabidopsis thaliana chromosome II BAC t17M13 genomic sequence.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Meyer S., Truernit E., Huemmer C., Besenbeck R., Stadler R., Sauer N.;
RT "AtSUC3; an unusual sucrose transporter from Arabidopsis.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004136; AAC32907.1; -;
DR EMBL; AJ289165; CAB92307.1; -;
SQ SEQUENCE 594 AA; 63972 MW; DF075A3601299DA0 CRC64;
```

```
Query Match 11.5%; Score 328.5; DB 10; Length 594;
Best Local Similarity 24.1%; Pred. No. 4.3e-15;
Matches 120; Conservative 72; Mismatches 174; Indels 131; Gaps 18;

QY 22 NLLTFGLEVCCLAAGITY-----VPPLLELVGEVEEKFTMTVVLGIPVLGVCVPLLG 73
DB 59 SLVTLVSTCTAAGVQFGWALQSLTTPYIQTILGISHAFSSFWLNCGPITGLVQVFGI 118
QY 74 ADHWRGRRRRPFTWALSGLTLLSLFIPRAGWLA--GLLCPD-----PRPLE 121
DB 119 WSDCKTSKYGRRRPFTLVGSMISIAVII---GFSADIGYLLGDSKEHCSTFKGTRTRA 175
QY 122 LALLILGVLLDFCGQCVFTPLEALLSDLFRPDHCRQAYSVAFMISLGGCLGCLYLLPAI 181
DB 176 AVVFIIGFWLLDLANNVTGQPARALLADL-SGPDORNTANAVFCULMMAIGNILGFSAGAS 234
QY 182 ----DWDTSALAPYLGTQECFLG-----LLTLIFLT-CVAAFTLLVAEEAALGPTPE 228
DB 235 GKQWEH-----FPFL-TSRACCAACGNLKAFLVAVVFLTICTLVITYFAKEIPFTSNKP 288
QY 229 AE-GLSAPSLPHCCPCARLARLNL-----253
DB 289 TRIQDSAPLLDD-----LQSKGLEHSLKNNGTANGIKYVERVEDTDEQFENSEHQ 340
QY 254 -----GALLPRLHQLCCMRPTLRLRVAELCSWMLMTFTLYTDFVGEGLYQGV 305
DB 341 DETYVDGPGSVLVNLLTSRLHLPAPMHSLVIVMALTWLSWFFFLFDDTMGSEVYHGD 400
QY 306 RAEPQTEARRHYDEGVRMGSLGLFLQCAISLVSFLVMDRLVQFGRVAVY-LASVAAFV 364
DB 401 TGD--SLHMELYDQGVREGALGLLNSVVLGISTSSFLPEPCORMGARVVWALSNETVFAC 458
QY 365 AGATCIS-----HSVAVVTASAAALGTGFTFSALQILPYTLASLY 403
DB 459 MAGTAVISMUSLSDKNGIEYIMRNETTRTAIV--FALLGFLPLAITYSPVFSVA-- 514.
QY 404 HREKQVFLPKYRGDTGG 420
DB 515 -----EVTADSGG 522

RESULT 15
ID 09FVL6 PRELIMINARY; PRT; 604 AA.
AC 09FVL6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE SUCROSE TRANSPORTER-LIKE PROTEIN.
GN SUT2.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20359188; PubMed=10899981;
RA Barker L., Kuhn C., Weise A., Schulz A., Gebhardt C., Hirner B.,
RA Hellmann H., Schulze W., Ward J.M., Frommer W.B.;
RT "SUT2, a putative sucrose sensor in sieve elements.";
RL Plant Cell 12:1153-1164(2000).
DR EMBL; AF166498; AAG12987.1; -;
SQ SEQUENCE 604 AA; 65082 MW; 3403349B1692B2AF CRC64;
```

```
Query Match 11.4%; Score 325.5; DB 10; Length 604;
Best Local Similarity 22.3%; Pred. No. 7e-15;
Matches 133; Conservative 89; Mismatches 210; Indels 165; Gaps 21;

QY 22 NLLTFGLEVCCLAAGITY-----VPPLLELVGEVEEKFTMTVVLGIPVLGVCVPLLG 73
```


Db 62 SLLTILSCVAAQVQFWALQSLTTPVIQTLGIEHAFSSFIWLCGPITGLVQPCVGI 121
QY 74 ASDHWRGRRPRPIWALSIGLISLFLIPRAGWLA--GLICPD-----PRPLE 121
Db 122 WSDKCHSKYGRRRPFIIGAVMISIAVII--GFSADIGYLLGDTKEHCSTFKGTRSR 178
QY 122 LALLILGVGLDFCGQVCFPLEALLSDLFDRPDHCRQAYSVYAFMISLGGCLGYLLPAI 181
Db 179 AIVFVGVFWMLDLANNVQGPALLADL--SGPDQNTANAVFCSMMAVGNILGFSAGAS 237
QY 182 D-----WDTSAAPYLGTOEEC-----LFGLLTILFT--CVAATLLVAEEAALGTE-P 228
Db 238 GQWHRW-----PFELTNRACCEPCGNLKAFLVAVVFLTCTLTLYFANEVPLSPKQYK 292
QY 229 ABGLSAPSL--SPH-----CCPCRARIAPENL----- 253
Db 293 RMSDSAPLLDSQONTGFDLSQSKRELQYRNSVANNESMGHVADNSPKNEEORPKDQGD 352
QY 254 -----GALLPRHLCCMPRTLRLFLVAELCSMMALMTFTLYTDFVGEGLYQVPR 307
Db 353 SFADSPGAVLVNLTSLRHLPPAMHSVLIVMALTWLSNFPFFLFDTDMMGREVYHGDPKG 412
QY 308 EPGTEARRHYDEGVMSGLGLFLOCAISLVFSLVMDRLVQREGTRAVYLA----- 358
Db 413 E--ADEVNAYNOGVREGAFGLLNSVVLGVSSFLIEPMCCKWIGSRVLVWAVSNFIVFCMA 470
QY 359 ----VAAPFVAAGATCLSHSAVV--TASAALTGTFSALQI-----LPYTLASLYHREK 407
Db 471 CTAIISVVSISANTQGVQHVIGATRSTQIAALVVFSLLLGIPLAVTYSVPFSITA----- 524
QY 408 QVFLPKYRGDTGGSSEDSLMTSELPKPGAPFNGHVAGGSGLLPPPPALCGASACD 467
Db 525 -----ELTADAG-----GGQGL----- 536
QY 468 VSVRVVGEPTEARVVPGRGICLDLAILDSAFLLSOVAPSLFMGSIVOLSQSVTAYM 524
Db 537 -----AIGVINLAIIVVQMVVSLGAGPWDALFGGNI--PAFALASLALAGIFAML 587

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2002, 07:56:35 ; Search time 13.31 Seconds
(without alignments)
1523.340 Million cell updates/sec

Title: US-09-605-783a-113

Perfect score: 2861
Sequence: 1 MVQRLWVSLRLHRKAQLL.....AIVFATQVDFKSLAKYSA 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	324.5	11.3	525	1 STP_SPIOL	Q03411 spinacia ol
2	142	5.0	544	1 YD74_SVNY3	P74168 synechocyst
3	125.5	4.4	399	1 TCR1_ECOLI	P02982 escherichia
4	121.5	4.2	503	1 PUR8_STRLP	P42670 streptomyce
5	119.5	4.2	491	1 AMFG_ECOLI	P36870 escherichia
6	118.5	4.1	496	1 SMVA_SALTY	P37594 salmonella
7	117	4.1	440	1 YHJE_ECOLI	P37643 escherichia
8	111.5	3.9	640	1 Y051_MYCTU	Q24723 nocardioides
9	111	3.9	473	1 PHDK_NOSK	P10880 mycobacteri
10	110.5	3.9	680	1 CAJA_HUMAN	Q03692 homo sapien
11	110	3.8	368	1 GALT_HUMAN	O60755 homo sapien
12	110	3.8	606	1 NQOC_THETH	Q56227 thermus aqu
13	109	3.8	477	1 YPUM_RHOCA	P26176 rhodobacter
14	107	3.7	419	1 CMLA_PSEAE	P32482 pseudomonas
15	105	3.7	448	1 YJ94_YEAST	P47159 saccharomyc
16	104	3.6	476	1 MELB_SALTY	P30878 salmonella
17	103	3.6	465	1 FTSM_MYCLE	O50186 mycobacteri
18	102.5	3.6	461	1 PUCC_RHOCA	P23462 rhodobacter
19	102	3.6	438	1 XHIA_ECOLI	P76350 escherichia
20	101.5	3.5	428	1 XYIO_BACSU	P42306 bacillus su
21	101	3.5	437	1 BRAZ_PSEAE	P25185 pseudomonas
22	101	3.5	471	1 MELB_ENTAE	O07366 enterobacte
23	101	3.5	471	1 MELB_KLEPN	O02581 klebsiella
24	100	3.5	404	1 Y4XM_RHISN	P55705 rhizobiella
25	99	3.5	481	1 LMRX_STRLN	P46104 streptomyce
26	98.5	3.4	613	1 NUOL_ECOLI	P36607 escherichia
27	98	3.4	379	1 CYB_SORCI	O79451 sorex ciner
28	98	3.4	496	1 GTR3_CHICK	P28568 gallus gall
29	97.5	3.4	385	1 P12R_BOVIN	P79393 bos taurus
30	97.5	3.4	433	1 DCUA_WOLSU	O34245 wolinnella s
31	97.5	3.4	1103	1 CYGD_HUMAN	Q02846 homo sapien
32	97	3.4	336	1 CYB_SORGA	O79996 sorex grana
33	97	3.4	336	1 CYB_SORHA	O79452 sorex hayde

RESULT 1

ID	STP_SPIOL	STANDARD;	PRT;	525 AA.
AC	Q03411;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	15-JUL-1998 (Rel. 36, Last annotation update)			
DE	SUCROSE TRANSPORT PROTEIN (SUCROSE PERMEASE) (SUCROSE-PROTON SYMPORTER).			
OS	Spinacia oleracea (Spinach).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Eukaryota; Magnoliophyta; eudicotyledons; core eudicots;			
OC	Caryophyllales; Caryophyllaceae; Chenopodiaceae; Spinacia.			
OX	NCBI_TaxID=3562;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Leaf;			
RA	MEDLINE=93099843; PubMed=1464305;			
RA	Riesmeier J.W., Willmitzer L., Frommer W.B.;			
RT	"Isolation and characterization of a sucrose carrier cDNA from spinach by functional expression in yeast.";			
RL	EMBO J. 11:4705-4713(1992)			
CC	-!- FUNCTION: RESPONSIBLE FOR THE TRANSPORT OF SUCROSE INTO THE CELL, WITH THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM). CAN ALSO TRANSPORT MALTOSE AT A LESSER RATE.			
CC	-!- PATHWAY: SUCROSE METABOLIC PATHWAY.			
CC	-!- SUBCELLULAR LOCATION: INNER MEMBRANE.			
CC	-!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).			
CC	EMBL; X67125; CAA47604.1; -			
DR	PIR; S28052; S28052.			
DR	InterPro; IPR003662; sub.trnsportr.			
DR	Pfam; PF000083; sugar.tr.1.			
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; FALSE_NEG.			
DR	PROSITE; PS0017; SUGAR_TRANSPORT_2; FALSE_NEG.			
KW	Transmembrane; Transport; Sugar transport; Symport; Glycoprotein. DOMAIN 1 40 CYTOPLASMIC (POTENTIAL).			
FT	TRANSMEM 41 60 POTENTIAL.			
FT	TRANSMEM 70 88 POTENTIAL.			
FT	TRANSMEM 108 127 POTENTIAL.			
FT	TRANSMEM 146 165 POTENTIAL.			
FT	TRANSMEM 185 204 POTENTIAL.			
FT	TRANSMEM 230 249 POTENTIAL.			
FT	TRANSMEM 296 315 POTENTIAL.			
FT	TRANSMEM 341 360 POTENTIAL.			
FT	TRANSMEM 373 392 POTENTIAL.			
FT	TRANSMEM 419 438 POTENTIAL.			
FT	TRANSMEM 441 461 POTENTIAL.			

```

FT TRANSMEM 490 509 POTENTIAL.
FT DOMAIN 510 525 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 525 AA; 54992 MW; 018347A4D2CC1CC6 CRC64;

Query Match 11.3%; Score 324.5; DB 1; Length 525;
Best Local Similarity 23.2%; Pred. No. 4.5e-16;
Matches 129; Conservative 91; Mismatches 196; Indels 141; Gaps 20;

QY 23 LITFGLVCLAGIY-----VPLLEVGVEEKEFTMWVLGICPVGLVGLVCPVLLGSA 74
Db 35 LKGLGLVASVAGVGFNALQLSLTPYVQLLGHPTWAAIYIWLCPISGMIVQLVGY 94

QY 75 SDHWGRYGRBRPFTWALSGLTLLSLFLIPRAGWLA--GLLCPDP-----RPLELALIL 127
Db 95 SDRCYRGRRRPFTIAGAALVAVAGLI---GFAADIGAASGDPTGNVAKPRATAVEV 151

QY 128 GYGLLDFFCGVCFTPLEALLSLDFR-DPDHCRQAYSVAEIMISLGGCLGYLLPAID--- 182
Db 152 GFWILDVANNTLOGPCRALLDMAAGSQTTRYANAFSFFMALGNIGYAGSRSRYT 211

QY 183 -----WDTSAIAPYLTQECIEFLGLT-LIFLTCVAATLLVAEEAALGPTEPAEGLSAPSL 237
Db 212 VEPFTKTAACDVYCANLKSCEFFISITLLIVLTILALSVMKQRIITIDEIQEEDLKNRN 271

QY 238 SPHCCPCRARLAFNIGALLPRLHOLCCRPRTLRRFLVAELCSWMLMFTFLFYDFVG 297
Db 272 SSGC-----ARLPF--FGOLIGALKDL-----PKPMILLLVLTALNMIAWFFLLFDTDMG 321

QY 298 EGLYGVPRAPPGTARRHYDEGRVMSGLGLFLOCAISLVFSLVMDRLVQRFQ--TRAVY 355
Db 322 KEVYGGT-----VGEGRLYDQGVHAGALGMINSVLVGNVSLSTEGRLRVVGGAKRLWG 375

QY 356 LASVAAPFVAGATCLSHSVAVVTASAAALGTFTSALQILPYTLASLYHREKQVFLPKYR 415
Db 376 IVNIIL-----AVCLAMTV-LVTKSAE-----HFRDSHHIM----- 405

QY 416 GDTGGASSEDLSMTSFLPGPKPGAPFPNGHVAGGSGGLLPPPPALCGASACDSVSVVVG 475
Db 406 -----GSAVPPPPPA--GVKGALAIFAVLG 429

QY 476 EPTEARV-VP-----GRGICLDLAIDSAFLLSQV-----AP 506
Db 430 IPLAITESIPPALASIFSASSGSGGLSLVNLAIIVPQMVFVTSQPDWAMFGGGLNP 489

QY 507 SLEWGSIVOLSQSTAY 523
Db 490 AFVVGAVAATAASAVLSF 506

```

```

RESULT 2
YD74_SYNY3
ID YD74_SYNY3 STANDARD; PRT; 544 AA.
AC P74168;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL SYMPORTER SLL1374.
GN SLL1374.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium

```

```

RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions." ;
RL DNA Res. 3:109-136(1996).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE SODIUM:GALACTOSIDE SYMPORTER FAMILY
CC (SGF).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC

```

```

EMBL; D90912; BAA18257.1; -
DR InterPro; IPR001927; Na_galact_symp.
DR Pfam; PF01236; Na_Galacto_symp; 1.
DR PROSITE; PS00872; NA_GALACTOSIDE_SYMP; FALSE_NEG.
DR Hypothetical protein; Transport; Transmembrane; Symport;
KW Complete proteome.
KW TRANSMEM 31 51 POTENTIAL.
FT TRANSMEM 52 72 POTENTIAL.
FT TRANSMEM 84 104 POTENTIAL.
FT TRANSMEM 116 136 POTENTIAL.
FT TRANSMEM 162 182 POTENTIAL.
FT TRANSMEM 191 211 POTENTIAL.
FT TRANSMEM 230 250 POTENTIAL.
FT TRANSMEM 257 277 POTENTIAL.
FT TRANSMEM 318 338 POTENTIAL.
FT TRANSMEM 356 376 POTENTIAL.
FT TRANSMEM 383 403 POTENTIAL.
FT TRANSMEM 407 427 POTENTIAL.
FT TRANSMEM 450 470 POTENTIAL.
FT TRANSMEM 501 521 POTENTIAL.
SQ SEQUENCE 544 AA; 60064 MW; C91D0EDFF32277EE CRC64;

```

```

Query Match 5.0%; Score 142; DB 1; Length 544;
Best Local Similarity 22.3%; Pred. No. 0.0049;
Matches 109; Conservative 60; Mismatches 179; Indels 126; Gaps 20;

QY 16 AQLLLVNLFLGLEVCLAAG-----ITYVPLLLLEV-GVEEFMTWLIGIPVL 63
Db 2 SOSLSAEKLHFTTKLAYGAGDFGPAITANILVFLYLLFLDTVDAGIPALAGSVLMIGKIF 61

QY 64 GLVCPVLLGASDHWGRYGRRRPFTWALSGLTLLSLFLIPRAGWLAGLCPDPRPLELA 123
Db 62 DAINDPIIGLLSDTRSRWGRRLPWM----LGGMIPFALFYTAQWLIPHFSDRLTNQWG 117

QY 124 LLILGVLLOFCQVCFT----PLEALLSDFRDPDHCROAYSVAEIMISLGGCLGYLLP 179
Db 118 LFYIYVAIA-MAFNLCYTTVNLPTALTPELTNYNE-RTRLNSFRFAFSGGSILSLIL 175

QY 180 AIDWDTSAIAPYLTQECIEFLGLL-TLIFLTCVAATLLVAEEAALGPTEPAEGLSAPSL 238
Db 176 YI-----LIAAGLPDRPQQQFGELGVMSLVLSISALLWSALRLOEKKEP---ILSPSL- 226

QY 239 PHCCPCRARLA-----FRNLGAL----- 256
Db 227 -----RRRLAPLLMAAGITLILLIAKSFNLLGSGDFYISFFLILLGLIWWGFGFTLR 280

QY 257 -----LPRHLQCCRPRTLRR-LFV--AELCSWMLMFTFLF 291
Db 281 DSAVEEHLQKLENSPSPGVVTENPLLLKQL--KIAFSNRAFLFVIGIYLSWLAVALQTASI 338

QY 292 YTDVGEGLYQGVPRAPPGTARRHYDEGRVMSGLGLFLOCAISLVFSLVMDRLVQRFGT 351
Db 339 LVYFVYS--WMGLNEQSGT-----IALAVQ-GTALVMLFWQALAQFLDK 381

QY 352 RAVY-LASVAAPFVAGATCLSHSVAVVTASAAALGTFTSALQILPYTL 399
Db 382 KVIYFLGSMVMWGAAGLWLVQPGVALLVTLAIFAGVGVSVAYLLIPWSM 431

```


RT The methyl viologen-resistance-encoding gene smvA of Salmonella
typhimurium.;
RL Gene 148:173-174(1994).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
CC AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D26057; BAA05055.1; -
DR StyGene; SG10384; smvA.
KW Transmembrane; Inner membrane
FT TRANSMEM 5 25 POTENTIAL.
FT TRANSMEM 44 64 POTENTIAL.
FT TRANSMEM 73 93 POTENTIAL.
FT TRANSMEM 96 116 POTENTIAL.
FT TRANSMEM 135 155 POTENTIAL.
FT TRANSMEM 158 178 POTENTIAL.
FT TRANSMEM 192 212 POTENTIAL.
FT TRANSMEM 220 240 POTENTIAL.
FT TRANSMEM 260 280 POTENTIAL.
FT TRANSMEM 299 319 POTENTIAL.
FT TRANSMEM 327 347 POTENTIAL.
FT TRANSMEM 357 377 POTENTIAL.
FT TRANSMEM 391 411 POTENTIAL.
FT TRANSMEM 470 490 POTENTIAL.
SQ SEQUENCE 496 AA; 52521 MW; 0F1B23C8FD27BCCD CRC64;

Query Match 4.1%; Score 118.5; DB 1; Length 496;
Best Local Similarity 22.5%; Pred. No. 0.21;
Matches 102; Conservative 44; Mismatches 143; Indels 165; Gaps 22;
QY 17 QLLVNLTLFGL-----EVCLAAGITVVPPLLELVG-----BEKEMT 54
DB 72 RLMLGGTTLFGLASLAFAAFSHTASMLIATRVLLAIGAAIVPATL-AGIRATFCEEKHN 130
QY 55 MVLGIGPVL---GLVCPVLGSADHWGRGRRRPFIWALSLGILL-----SLFLI-- 103
DB 131 MALGVAAVSGGRAFPVLIG-----GILLHFYVGSVFLINY 168
QY 104 PRAGWLAGLLCP-DPR-----PLEL---ALLILGVLLDFCGQ-----VCPT 141
DB 169 PIVLVVMGLTARYDPRQGRDQPLNLGHVVMILITAILLVYSAKTALKHLSLWISVT 228
QY 142 PLE-ALLSLDFRDPD-----PLEL---ALLILGVLLDFCGQ-----VCPT 141
DB 229 LITGALLGLFRTQATSRPMIDMRFLFTHRIILSGVVMAMTITLVGFELMAQELQF 288
QY 156 -HCRQAYSAYAFMI-----SIGGCL-----GYLLPAIDMDTSLAPYL 192
DB 289 VHGLSPYEAGVFMPLVWVASGFSPTAGVLVSRGLRLVATGMALSALSFYGLAMTDF- 347
QY 193 GTQECLFGLLLIFLTCVAATLLVAEEAALGFTPEAE-GLSAPLSLPHCCPCRARLAFR 251
DB 348 STQQWQAWGLMALLGFS--AASALLASTSAIMAAAPAEKAAAGAIETMAYELGAGLGTA 405
QY 252 NLGALLPRHLQCCMRPTLRFLVAELCSWMA-LMTFTFLYTFDVFEGLEGYGVRAEPG 310
DB 406 IFGLLSRSFSASIRLPAGLAEQAEIARASSNGEAVQLANSYPTTQGGKY----- 456
QY 311 TEARRH---YDGEVRMGLG-LFLOCAISLVFSL 340
DB 457 LTAARHAFIWSHVALSSAGSMLLLLAVGMWFSL 490

RESULT 7
YHJE_ECOLI
ID YHJE_ECOLI STANDARD; PRT; 440 AA.
AC P37643;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL METABOLITE TRANSPORT PROTEIN YHJE.
GN YHJE OR B3523.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=94316500; PubMed=8041620;
RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
RT region from 76.0 to 81.5 minutes";
RL Nucleic Acids Res. 22:2576-2586(1994).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. STRONG, TO
CC H.INFLUENZAE HI0281 AND HI0418.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U00039; AAB18499.1; -
DR EMBL; AE000429; AAC76548.1; -
DR EcoGene; EG12249; yHJE.
DR InterPro; IPR003662; sub_trnsprtr.
DR Pfam; PF00083; sugar_tr.1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.
KW Hypothetical protein; transport; Transmembrane; Inner membrane;
KW Complete proteome.
FT TRANSMEM 35 55 POTENTIAL.
FT TRANSMEM 67 87 POTENTIAL.
FT TRANSMEM 109 129 POTENTIAL.
FT TRANSMEM 130 150 POTENTIAL.
FT TRANSMEM 168 188 POTENTIAL.
FT TRANSMEM 193 213 POTENTIAL.
FT TRANSMEM 249 269 POTENTIAL.
FT TRANSMEM 290 310 POTENTIAL.
FT TRANSMEM 321 341 POTENTIAL.
FT TRANSMEM 346 366 POTENTIAL.
FT TRANSMEM 385 405 POTENTIAL.
FT TRANSMEM 411 431 POTENTIAL.
SQ SEQUENCE 440 AA; 47208 MW; 3A23302A9DCBE5BE CRC64;
Query Match 4.1%; Score 117; DB 1; Length 440;
Best Local Similarity 28.4%; Pred. No. 0.24;
Matches 48; Conservative 29; Mismatches 54; Indels 38; Gaps 10;
QY 53 MFMVLGIGPVLGVLCVPLLGASDHWGRGRRRPFIWALSLGILLFLI-PRAGWLAG 111
DB 295 MMVAVTGFQ-----VMVPVAGLIAD---AFGRKSMVITITLILFAFENPLIG---- 341
QY 112 LCLCPDRP-LELALLILGVLLDFCGQVCFPLEALLSLDFRDPDHCQAVSVYAFMISL 170
DB 342 -----SGNPLIVPAFLLEGLSLMG-----LTFGPMGALLPELP--PTEVR--YTGFSEFNV 389
QY 171 GCCLGYLLPAIDMDTSLAPYLGTQECLFGLLLT-IFLTCVVAATLLVA 218
DB 390 ASILG-----ASVAPYIAAWLQTNVGLGAVGLVLAAMAGTLTIA 428

```
RESULT 8
Y051_MYCTU
ID Y051_MYCTU STANDARD; PRT; 640 AA.
AC Q10880; OS3628;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHEICAL 65.6 KDA PROTEIN RV0083.
GN RV0083 OR MT0090 OR MTCY251.01 OR MTV030.27.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
Badoock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -!- SIMILARITY: TO POLYPEPTIDE 5 OF THE NADH-UBIQUINOL OXIDOREDUCTASE
OF CHLOROPLASTS OR MITOCHONDRIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL021428; CAAL6264.1; -
CC EMBL; AE006920; AAK44315.1; -
CC TIGR; MT0090; -
CC TubercuList; Rv0083; -
CC InterPro; IPR001750; Oxidored_q1.
CC Pfam; PF00361; oxidored_q1; 1.
CC Hypothetical protein; Oxidoreductase; Transmembrane;
KW Complete proteome.
FT TRANSMEM 8 28 POTENTIAL.
FT TRANSMEM 52 72 POTENTIAL.
FT TRANSMEM 90 110 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 179 199 POTENTIAL.
FT TRANSMEM 208 228 POTENTIAL.
FT TRANSMEM 241 261 POTENTIAL.
FT TRANSMEM 277 297 POTENTIAL.
FT TRANSMEM 298 318 POTENTIAL.
FT TRANSMEM 352 372 POTENTIAL.
FT TRANSMEM 391 411 POTENTIAL.
FT TRANSMEM 446 466 POTENTIAL.
```

```
FT TRANSMEM 497 517 POTENTIAL.
FT TRANSMEM 619 639 POTENTIAL.
SQ SEQUENCE 640 AA; 65627 MW; 6254C289DBD108A8 CRC64;

Query Match 3.9%; Score 111.5; DB 1; Length 640;
Best Local Similarity 22.3%; Pred. No. 0.86; 222; Indels 155; Gaps
Matches 129; Conservative 73; Mismatches 73;

QY 16 AOLLLV---NLLTFGL--EVCALAAGITV-----PPLLLEVGVEEKFMTWVLGICPVLG 64
Dy 93 AAMLLVPAAGSVTFLLAELMELMASLILVSEHARPOVRSGAGLWYAVMTQLGFTAILVG 152
QY 65 LVCVPLLGASDHRGRYRRRPFIALSLGILLSLF-----LIPRAGWLGL 112
Dy 153 LVVLAAG--GSDRFAG-----LGAVCDGVRAAVFMTLVFGSGKAGLVPFLHAWLPRA 203
QY 113 LCPDPRP-----LELALLILGV-GLLDFCGQVCFPLEALLSDFDPDHCQAYSVYAFM 167
Dy 204 HPEAPSPVSALMSAAMNIGIYVFDLQL-----LGGPRWGLA-----L 246
QY 168 ISLGS---CLGVLLPAIDWDTSALAPYLGTOECLFGLTFLTCVAATLLVADEAALG 224
Dy 247 LAVGTSALYGVQASVAADLKRLAYSTTEN--MGLITL----ALGAATLFDATGAYG 299
QY 225 PTEAEGLSAPSLSPHCCPCRARLAFRNLGALLPR-----LHQLCCMRPRTLRLRV 276
Dy 300 PASTAAAAAAMLMIAHA--AFKSLAFMAAGSVLAATGLRDLGLGLARRMPAT----- 351
QY 277 AELCSWMLMTFTLFY-----TDFVGEGLYGVPRAPETGTEARRHYDEGVR 322
Dy 352 -----TVFFGVAALGACGLPLGAGFVSEMLLVQSLIHAAPGDPITVALTTPLA 399
QY 323 MGSGLFLQCAISLVES-----LVMDRLVQRFGRVAVLASVAAPFVAAAGATCLSHVA 376
Dy 400 VGVALATGLSVAAMTKAFGICFLAPRSTQAEAREAPASRAGMAITAAGA-CLVLAVA 458
QY 377 --VVTASALGFTFSALQILPYT-LASLYHREKOVFLPKYRGDTGGASSEDLSMTSLP 433
Dy 459 PLLVAPMVRRAAATLPAAQAVKFTGLGAV-----VRLPAMSG-----SIAP 499
QY 434 GPKGAPFPNPHVAGGSGLLP-----PEPA-----LCGASACDVSVRVVVGEPTAR 481
Dy 500 GVIAAAIL-----AAALAVAVLARWFRRRPAPARLPLWACGAA--DLTVRMQYATSPAE 553
QY 482 VVPGIGICLDLAIDLSAFLSQVAPSLFMSIVQLSQSV 520
Dy 554 --PLQRFVGDVLRPDTDIEVTHTAESRYMAERITYRTAV 590

RESULT 9
PHDK_NOCCK STANDARD; PRT; 473 AA.
AC O24723;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PROBABLE 1-HYDROXY-2-NAPHTHOATE TRANSPORTER.
OS Nocardioides sp. (strain KP7).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Propionibacterineae; Nocardioidaceae; Nocardioides.
OX NCBI_TaxID=35761;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=97474276; PubMed=9335300;
RA Iwabuchi T., Harayama S.;
RT "Biochemical and genetic characterization of 2-carboxybenzaldehyde
dehydrogenase, an enzyme involved in phenanthrene degradation by
Nocardioides sp. strain KP7.";
RL J. Bacteriol. 179:6488-6494(1997).
CC -!- FUNCTION: PROBABLY INVOLVED IN THE TRANSPORT OF 1-HYDROXY-2-
NAPHTHOATE.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
```


CC (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB000735; BAA23264.1; .
DR InterPro: IPR003662; sub_trnsprtr.
DR Pfam: PF00083; sugat_tr; 1.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.
KW Transport; Transmembrane; Inner membrane.
FT TRANSMEM 26 46 POTENTIAL.
FT TRANSMEM 60 80 POTENTIAL.
FT TRANSMEM 92 112 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 154 174 POTENTIAL.
FT TRANSMEM 181 201 POTENTIAL.
FT TRANSMEM 264 284 POTENTIAL.
FT TRANSMEM 305 325 POTENTIAL.
FT TRANSMEM 331 351 POTENTIAL.
FT TRANSMEM 355 375 POTENTIAL.
FT TRANSMEM 416 436 POTENTIAL.
SQ SEQUENCE 473 AA; 49109 MW; D6D765D376260D8A CRC64;

Query Match 3.9%; Score 111; DB 1; Length 473;

Best Local Similarity 24.3%; Pred. No. 0.68;

Matches 114; Conservative 66; Mismatches 180; Indels 110; Gaps 27;

QY 12 RHRKAQLLVNLTFF--GLEVCVLAAGITVYPLLELVGVVEEFMTMVLGIGVPLGVCP 69
DB 22 RQRRIIVLVFIMVADGMDITLASHL--FPVIRDWGVPSAVTLVSLGVVVAIGAL 79
QY 70 LIGSASDHRVGRYRRRPFINALSGLISLIPRAG-WLAGLLCPDRPRLEALLIIG 128
DB 80 VSGPVADRW-GRKG-----VTVVGFV--LFCATAGLGLTG-----DIHSFAALRIISC 125
QY 129 VGLLDFCGVCPTLEALLSLDFRPDCHRC-QAYSVYAFMISLGCGLGVLPAIDWDTSA 187
DB 126 FGL---GAVMPVAL-TIVADW--PKARRAQWVSIAFAGVGVSTIGAYLAA-----A 172
QY 188 LAPYLGTQECOLF-GLLTLIFITCVAAATLLVAEEA-----ALGPTPEAG 231
DB 173 VIPTLGMQVWVLIAGLAPLIILPFFVA--LYPEPAISVVRGIPPEARIRSAALVAPDRD 230
QY 232 LSAPSLPHCCPCRRARLAFRNGALLPRHLQCCRMPTRLRLFVAELCSWMAIMFTFLF 291
DB 231 IAGVDLT-----RAGL---TLGAGEVR-----AKALFAEILCRPLLGVTLLIW 270
QY 292 YTDFFVCEG---LYQGVF---RAEPGTERRHYDEGVEMSGSLGLFLQCAISLVESLYMD 343
DB 271 GVFFVQSGSLVLQIMPLMQAPAGLST-----VESGLIVAMYGW-----ALIGOLTIA 322
QY 344 RLVQRFGRVAVLASVAPFVAAAGATCLSHSVAVVATASALTGTF-----SALQI-L 395
DB 323 FILKRF-DRFIALAAFIWSV-----VGLLIVAFAGTGFVGFYFTLLFAIGLSL 371
QY 396 PTLASLYHREKQVLPKYR-----GDTGASSEDSLMTSFLPGPKPGAPF 441
DB 372 PATAAMQSVTTLAYEEFRATGMGSAGFAGRLGLTLYGALGGLTIGAGF 421

RESULT 10

CAIA_HUMAN

ID CAIA_HUMAN STANDARD; PRT; 680 AA.

AC Q03692;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE COLLAGEN ALPHA 1(X) CHAIN PRECURSOR.
GN COL10A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92109659; PubMed=1764025;
RA Thomas J.T., Cresswell C.J., Rash B., Nicolai H., Jones T.,
RA Solomon E., Grant M.E., Boot-Handford R.P.;
RT "The human collagen X gene. Complete primary translated sequence and
RT chromosomal localization.";
RL Biochem. J. 280:617-623(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93012005; PubMed=1397333;
RA Reichenberger E., Beier F., Luvalle P., Olsen B.R., von der Mark K.,
RA Bertling W.M.;
RT "Genomic organization and full-length cDNA sequence of human collagen
RT X.";
RL FEBS Lett. 311:305-310(1992).
RN [3]
RP SEQUENCE FROM N.A.
RA Beier F., Lammi M.B., von der Mark K.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 52-680 FROM N.A.
RX MEDLINE=92267014; PubMed=1587271;
RA Apte S.S., Seldin M.F., Hayashi M., Olsen B.R.;
RT "Cloning of the human and mouse type X collagen genes and mapping of
RT the mouse type X collagen gene to chromosome 10.";
RL Eur. J. Biochem. 206:217-224(1992).
RN [5]
RP SEQUENCE OF 561-666 FROM N.A.
RX MEDLINE=91243838; PubMed=2037056;
RA Apte S., Mattei M.-G., Olsen B.R.;
RT "Cloning of human alpha 1(X) collagen DNA and localization of the
RT COL10A1 gene to the q21-q22 region of human chromosome 6.";
RL FEBS Lett. 282:393-396(1991).
RN [6]
RP SEQUENCE OF 547-655 FROM N.A.
RX MEDLINE=92077285; PubMed=1743401;
RA Reichenberger E., Aigner T., von der Mark K., Stoeb H., Bertling W.;
RT "In situ hybridization studies on the expression of type X collagen
RT in fetal human cartilage.";
RL Dev. Biol. 148:562-572(1991).
RN [7]
RP REVIEW ON VARIANTS.
RX MEDLINE=97255959; PubMed=9101290;
RA Kuivaniemi H., Tromp G., Prockop D.J.;
RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-
RT associated collagen (type IX), and network-forming collagen (type X)
RT cause a spectrum of diseases of bone, cartilage, and blood vessels.";
RL Hum. Mutat. 9:300-315(1997).
RN [8]
RP VARIANTS SMCD ASP-598 AND PRO-614.
RX MEDLINE=94136476; PubMed=8304336;
RA Wallis G.A., Rash B., Sweetman W.A., Thomas J.T., Super M., Evans G.,
RA Grant M.E., Boot-Handford R.P.;
RT "Amino acid substitutions of conserved residues in the
RT carboxyl-terminal domain of the alpha 1(X) chain of type X collagen
RT occur in two unrelated families with metaphyseal chondrodysplasia
RT type Schmid.";
RL Am. J. Hum. Genet. 54:169-178(1994).
RN [9]
RP VARIANT SMCD ARG-591.
RX MEDLINE=9427470; PubMed=8004099;
RA McIntosh I., Abbott M.H., Warman M.L., Olsen B.R., Francomano C.A.;
RT "Additional mutations of type X collagen confirm COL10A1 as the
RT Schmid metaphyseal chondrodysplasia locus.";

RL Hum. Mol. Genet. 3:303-307(1994).
RN [10]
RP VARIANT SMCD VAL-618.
RX MEDLINE-95181449; PubMed-7876225;
RA Chan D., Cole W.G., Rogers J.G., Bateman J.F.;
RT "Type X collagen multimer assembly in vitro is prevented by a Gly618
RT to Val mutation in the alpha 1(X) NC1 domain resulting in Schmid
RT metaphyseal chondrodysplasia.";
RL J. Biol. Chem. 270:4558-4562(1995).
RN [11]
RP VARIANTS SMCD ARG-545; GLU-595; HIS-597; LYS-617; ARG-644 AND GLY-648.
RX MEDLINE-95331767; PubMed-7607655;
RA Bonaventure J., Chaminade F., Maroteaux P.;
RT "Mutations in three subdomains of the carboxy-terminal region of
RT collagen type X account for most of the Schmid metaphyseal
RT dysplasias.";
RL Hum. Genet. 96:58-64(1995).
RN [12]
RP VARIANT SMCD PRO-600.
RX MEDLINE-96375754; PubMed-8782043;
RA Wallis G.A., Rash B., Sykes B., Bonaventure J., Maroteaux P.,
RA Zabel B., Wynne-Davies R., Grant M.E., Boot-Hanford R.P.;
RT "Mutations within the gene encoding the alpha 1 (X) chain of type X
RT collagen (COL10A1) cause metaphyseal chondrodysplasia type Schmid but
RT not several other forms of metaphyseal chondrodysplasia.";
RL J. Med. Genet. 33:450-457(1996).
RN [13]
RP VARIANTS SMCD GLU-18 AND ARG-18.
RX MEDLINE-97220591; PubMed-9067753;
RA Ikegawa S., Nakamura K., Nagano A., Haga N., Nakamura Y.;
RT "Mutations in the N-terminal globular domain of the type X collagen
RT gene (COL10A1) in patients with Schmid metaphyseal
RT chondrodysplasia.";
RL Hum. Mutat. 9:131-135(1997).
RN [14]
RP VARIANTS SMD GLU-595.
RX MEDLINE-99057503; PubMed-9837818;
RA Ikegawa S., Nishimura G., Nagai T., Hasegawa T., Ohashi H.,
RA Nakamura Y.;
RT "Mutation of the type X collagen gene 'COL10A1' causes
RT spondylometaphyseal dysplasia.";
RL Am. J. Hum. Genet. 63:1659-1662(1998).
RN [15]
RP VARIANT SMCD CYS-597.
RX MEDLINE-99069781; PubMed-9852679;
RA Sawai H., Ida A., Nakata Y., Koyama K.;
RA "Novel missense mutation resulting in the substitution of tyrosine by
RT cysteine at codon 597 of the type X collagen gene associated with
RT Schmid metaphyseal chondrodysplasia.";
RL J. Hum. Genet. 43:259-261(1998).
CC -1- FUNCTION: TYPE X COLLAGEN IS A PRODUCT OF HYPERTROPHIC
CC CHONDROCYTES AND HAS BEEN LOCALIZED TO PRESUMPTIVE
CC MINERALIZATION ZONES OF HYALINE CARTILAGE.
CC -1- SUBUNIT: HOMOTRIMER.
CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -1- DISEASE: DEFECTS IN COL10A1 ARE THE CAUSE OF SCHMID TYPE
CC METAPHYSEAL CHONDRODYSPLASIA (SMCD), A DOMINANTLY INHERITED
CC DISORDER OF THE OSSEOUS SKELETON. THE CARDINAL FEATURES OF THE
CC PHENOTYPE ARE MILD SHORT STATURE, COXA VARA AND A WADDLING GAIT.
CC RADIOGRAPHY USUALLY SHOWS SCLEROSIS OF THE RIBS, FLARING OF THE
CC METAPHYSES, AND A WIDE IRREGULAR GROWTH PLATE, ESPECIALLY OF THE
CC KNEES.
CC -1- DISEASE: DEFECTS IN COL10A1 ARE THE CAUSE OF SPONDYLOMETAPHYSEAL
CC DYSPLASIA (SMD). SMD COMPRISES A HETEROGENEOUS GROUP OF HERITABLE
CC SKELETAL DYSPLASIAS CHARACTERIZED BY MODIFICATIONS OF THE
CC VERTEBRAL BODIES OF THE SPINE AND METAPHYSES OF THE TUBULAR BONES.
CC -1- SIMILARITY: STRONG, TO ALPHA 1 AND 2 TYPE VIII COLLAGENS.
CC -1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce>,
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X60382; CAA42933.1; -
CC EMBL; X65120; CAA46236.1; -
CC EMBL; X9568; CAA67178.1; -
CC EMBL; X68531; AAC60615.1; -
CC EMBL; X58879; CAA41686.1; -
CC EMBL; M74050; AAA61221.1; -
CC EMBL; X72579; CAA51170.1; -
CC EMBL; X72580; CAA51170.1; JOINED.
CC PIR; S15826; S15826.
CC PIR; S30086; S30086.
CC PIR; A43901; A43901.
CC PIR; S18249; S18249.
CC PIR; S21856; S21856.
CC PIR; S26396; S26396.
CC MIM; 120110; -
CC MIM; 156500; -
CC MIM; 184250; -
CC InterPro: IPR001073; Clq.
CC InterPro: IPR000087; Collagen.
CC Pfam: PF00386; Clq; 1.
CC Pfam: PF01391; Collagen; 7.
CC PRINTS: PR00007; COMPLEMENTC1Q.
CC SMART; SM00110; Clq; 1.
CC PROSITE; PS01113; C1Q; 1.
CC Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
CC Cartilage; Collagen; Signal; Disease mutation; Polymorphism.
CC SIGNAL; 1 18
CC CHAIN; 19 680
CC DOMAIN; 19 56
CC DOMAIN; 57 519
CC DOMAIN; 520 680
CC DOMAIN; 545 680
CC VARIANT; 18 18
CC VARIANT; 18 18
CC VARIANT; 545 545
CC VARIANT; 591 591
CC VARIANT; 595 595
CC VARIANT; 597 597
CC VARIANT; 597 597
CC
CC Query Match 3.9%; Score 110.5; DB 1; Length 680;
CC Best Local Similarity 30.7%; Pred. No. 1.1;
CC Matches 43; Conservative 11; Mismatches 51; Indels 35; Caps 7;
QY 412 PKYRGDTG--GASSEDLSMTSFLPGP--KPGAPPNGHVGAGSGLLPPPPALCGASACD 467
DB 463 PGSGKDPGSGPPGPGAGIATKGLNGTGPFGPPGPRGHSGEPPG--LPGPP----- 510
QY 468 VSVRVVVGTEARVVPVGRICIDLDLAILDSAFLLSQVAPSLFVGSIVOLQSQSVTAYMVSA 527
DB 511 -----GPPGP----PGQ-----AVMPEGFIKAGORPSLSGTPLVSAQGVGTGMPVSA 553
QY 528 AGLGLVAIY--FATQVVFEDK 545
DB 554 FTVILSKAYPAIGTPIPFEDK 573
RESULT 11
GALT_HUMAN
ID GALT_HUMAN STANDARD; PRT; 368 AA.
AC O60755;

QY 139 CFTPEALLSDLPDPCHCRQAYSVVAFMISLGCLGYLLPAIDWDTLSALAPYLGTQEEC 198
Db 74 CCVPFOATI-----YTLD--W-----88
QY 199 LFGLLT-----LIFLTCVAATLLVAERAAALGPTPEAGLSAPSLSPHCCPCRRARLAFRN 252
Db 89 LFGALCVKAVHLLIYLTMYASFT-----LAAVSVDY-----LAVR- 125
QY 253 LGALLPRLHOLCCMRPTLRLRFAELCSMMALMTFTLFTDFVGSGLYQGVPRAPGTE 312
Db 126 -----HPLRGRALTPRNARAAGLVWLLAALFSAPYLSYGTGVRYGALELCVPAME 177
QY 313 ARHYDEGVRMGSLGLFLQCA--ISLVFSLVMDRLVORFG-----TRAVLASVAA 361
Db 178 DARRALDVATFAAGYLLPVAVVSLAYGRTLRFLNRAVGPAGAAAEARRRATGRAGRAM 237
QY 362 FPAVA-----GATCLSHSVAVVTASAAALTGFTFSALQILPYTTLASLY 403
Db 238 LAVAALYALCWGPHHALILCFWYGRFAFSPATYACRLASHCLAYANSCLNPLVYALASRH 297
QY 404 -----HREKQVLPKYRGDTGGASSEDLSMTSFLPCPKGAPFNGHVG 447
Db 298 FRARFRLPCGRRRRHARRA-LRRVRPASSG-----PPGCGDARPSGRLL 344
QY 448 AGSGGLPPP 457
Db 345 AGG-GQPEP 353
RESULT 12
NOOC_THETH STANDARD; PRT; 606 AA.
AC Q56227;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 12 (EC 1.6.5.3) (NADH
DE DHYDROGENASE 1, CHAIN 12) (NDH-1, CHAIN 12).
GN NOO12.
OS Thermus aquaticus (subsp. thermophilus).
OC Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB8 / ATCC 27634;
RX MEDLINE=97172490; PubMed=9020134;
RA Yano T., Chu S.S., Sled V.D., Ohnishi T., Yagi T.;
RT "The proton-translocating NADH-quinone oxidoreductase (NDH-1) of
thermophilic bacterium Thermus thermophilus HB-8. Complete DNA
sequence of the gene cluster and thermostable properties of the
expressed NQO2 subunit";
RL J. Biol. Chem. 272:4201-4211(1997).
CC -1- SUBUNIT: COMPOSED OF 14 DIFFERENT SUBUNITS. SUBUNITS NQO7-14
CC CONSTITUTE THE MEMBRANE SECTOR OF THE COMPLEX.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: TO POLYPEPTIDE 5 OF THE NADH-UBIQUINOL OXIDOREDUCTASE
OF CHLOROPLASTS OR MITOCHONDRIA.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL: U52917; AAA97949.1;
DR InterPro: IPR003916; NADhub_oxdrdctse5.
DR InterPro: IPR001750; Oxidored_q1.
DR InterPro: IPR001516; Oxidored_q1_N.
DR pfam: PF00361; oxidored_q1; 1.

DR pfam: PF00662; oxidored_q1_N; 1.
KW Oxidoreductase; NAD; Ubiquinone; Transmembrane.
FT TRANSMEM 1 21 POTENTIAL.
FT TRANSMEM 29 49 POTENTIAL.
FT TRANSMEM 75 95 POTENTIAL.
FT TRANSMEM 104 124 POTENTIAL.
FT TRANSMEM 126 146 POTENTIAL.
FT TRANSMEM 165 185 POTENTIAL.
FT TRANSMEM 206 226 POTENTIAL.
FT TRANSMEM 234 254 POTENTIAL.
FT TRANSMEM 266 286 POTENTIAL.
FT TRANSMEM 303 323 POTENTIAL.
FT TRANSMEM 345 364 POTENTIAL.
FT TRANSMEM 364 384 POTENTIAL.
FT TRANSMEM 397 417 POTENTIAL.
FT TRANSMEM 441 461 POTENTIAL.
FT TRANSMEM 492 512 POTENTIAL.
FT TRANSMEM 584 604 POTENTIAL.
SQ SEQUENCE 606 AA; 65141 MW; 47318BF2665295C CRC64;
Query Match 3.8%; Score 110; DB 1; Length 606;
Best Local Similarity 21.4%; Pred. No. 1;
Matches 126; Conservative 69; Mismatches 198; Indels 196; Caps 27
QY 62 VLGLVCVPLGSASDHWRGGR--RRPFIWALSGLILLSLFLIP-----RAGW 108
Db 3 LLGTILLPLLGFA---LLGLFKRMREPLPGVILASGLVCLASFLLGAGLLSGGARQAEW 59
QY 109 LAGL---LCPDPRPLEALLILGVGLLDFCGQVCFTPLEALLSDLPDPCHCRQAYSVY- 164
Db 60 LPTGFSLLDNLGSGFMILLIVTGVGLIHVYAGY-----MGDPGYSR--FPAYF 108
QY 165 ---AFMISLGGCLGYLLPAIDWDTLSALAPYLGTQEECLFG-----201
Db 109 NLFIAMMLTLVADSYPMVFIGWEGVGLASFL-----LICFPYKNQYADSAKAFIVN 162
QY 202 -----LITLIELTCVAATLLVAE--EALGPTPEAGLSAPSLSPHCCPCRRARLAFRN 252
Db 163 RIGDLGFMGLGMAILWALYGLTSLISELKEAMEGP-----LKNPDL-----LALAG 206
QY 253 LGALLPRLHOLCCMRPTLRLRFAELCSMMALMTFTLFTDFVGEV-----LYQGVPRAE 308
Db 207 L-----LFLGAVGKSAQIPLMWLPLDAM 230
QY 309 PGTEARRHDEGVRMGSLGLFLQCAISLVFSLVMDRLVORFGTRAVVTLASVAAFFVAA-G 367
Db 231 AGTPVSALIIHAATMTVGTAGYILARSFSLYSVLFD---VSTAIYVGLLTAAYCALSAFG 287
QY 368 ATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVF-----410
Db 288 QTDIKKIVAYSTISQ--LGYMFLAAGVAGYVVA-LPHVETHAFPKALLFLASGSVHALG 344
QY 411 -----LPKYRGDT-----GGASSEDLSMTSFLPCPKGAPFPN 443
Db 345 GEODVRKMGGLWKHLQTRMHALIGALALGGLPLLSGFSKDAIATLATIYPPFGVGFG-- 402
QY 444 GHVGAGSGGLPPPPALCGASACDVSVRVVVG-----PTEARV---PGRGICLDIAI 494
Db 403 -YGA----LL--VAVLTAMYARNRWFVLVFLGEERHHHPHPPVLMNPHNLLALG-SV 454
QY 495 LDSAFLLSQVAPSLFMGSIYVOLQSOTAYMVS--AAGLGLVAIYATQVV 542
Db 455 LAGYIALPHLPNVLPEFLKPAALAEVDAHHLSLGAENGLSALSAVALL 503
RESULT 13
YPOM_RHOCA STANDARD; PRT; 477 AA.
ID YPOM_RHOCA
AC P26176;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-MAY-1992 (Rel. 22, Last annotation update)

DE HYPOTHETICAL 50.4 KDA PROTEIN IN PUHA-BCHM INTERGENIC REGION (ORF477)
DE (PROTEIN F1696).
OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84259352; PubMed=6744416;
RA Youvan D.C., Bylina E.J., Alberti M., Begusich H., Hearst J.E.;
RT "Nucleotide and deduced polypeptide sequences of the photosynthetic
RT reaction-center, B870 antenna, and flanking polypeptides from R.
RT capsulata.";
RL Cell 37:949-957(1984).
CC -!- SIMILARITY: TO RHODOBACTER CAPSULATUS PUCC AND TO
CC BACTERIOCHLOROPHYLL SYNTHASE 44.5 KDA CHAIN.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenseseis@sib.ch).

CC ENBL: Z11165; CAA77521.1; .-
DR EMBL: K01183; -. NOT_ANNOTATED_CDS.
DR PIR: B28988, B28988.
KW Photosynthesis; Hypothetical protein.
SQ SEQUENCE 477 AA; 50372 MW; 09AGA50AECDD4D998 CRC64;

Query Match 3.8%; Score 109; DB 1; Length 477;
Best Local Similarity 21.9%; Pred No. 0.96;
Matches 122; Conservative 80; Mismatches 178; Indels 176; Gaps 29;

QY 17 QLLLVNL--LTFGLEVCLAGITVPPLLLEVGVEEKFTMTVLGIGPVLGLVCVPLLGSA 74
DB :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
DB 30 QLRLSLFQVSVMQAQVLLLG-TLNRMVILELGVPAALVVAAMISI-PVLVAPFRAILGHR 87
QY 75 SDHWRYGY-RRRPFWALSIGILLSFLIPRA-----GW---LAGLLCPDP 117
DB :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
DB 88 SDTRSLGWNKVRPYLMFGLSNMGGAALMPFSILLSGDQTGMGPWAAGRAFAG----- 141
QY 118 RPLEALLILGVGLLDPCGOVCTPTLEALLSDLPDPHCRQAYSVA-FMISLG---GC 173
DB :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
DB 142 ----VAFLMAGVM----HMTQTAGLAADRATEETRPQVALLVFMFLGMSAVI 192
QY 174 LGYLLPAIDWDTSAPIYLGTQEELGLTLPLTCVAATLLVAEBAALGPTPEAGLS 233
DB :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
DB 193 VGWLLR--DFDOIPL---IRWOGC--GAMTLV-LNVIA---LWKOEVMRPMTKAER-RE 239
QY 234 APSLSPHCCPCRARLARPNLGRHLPRHQLCRCMRPTLRRLFVAELCSWMALMTETLFYT 293
DB :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
DB 240 AP-----RQSFEAWGLL-----RAETGALRLAT----- 264
QY 294 DFVGEGLYQGVPRAEPTGARHYDEGVRMGSGLGLFOCALISLVSLMDRLVQRFGTRA 353
DB :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
DB 265 -----VMVGTAFASMQ-----DVLLPEYGGQV 286
QY 354 VYLASVAAPFAAGATCLSHSVAVVTASAALTGTFESALOILPYTLASLVHREKVFLPK 413
DB :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
DB 287 LGLKVGQTTWLTA-----WAFGALVGFIWSARRL---SQGAVAHVAA----- 327
QY 414 YRGDTGASSSEDSLTSFLPGPKPGAPPFNHGAG-GSGLLPPPPLACASACDSVRY 472
DB :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
DB 328 -RGILLYGIVAFTVLSPFLPSGSK-VLFFASAMGIGUGSMF-----GIATLTVMVV 377
QY 473 VW-----GEPTEARVPVGRGICLDIALDSAPLLSQVAPSFLPMGSIVOLQS 519
DB :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
DB 378 VVRGASGIALGANGAQATAA----GLAVFIGGATRD----LVAAAAAGVLSLHSPALG 430
QY 520 VTAIMVSAAGLGLVAI 535

FT DOMAIN 163 169 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 170 187 6 (POTENTIAL).
FT DOMAIN 188 214 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 215 235 7 (POTENTIAL).
FT DOMAIN 236 245 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 246 266 8 (POTENTIAL).
FT DOMAIN 267 279 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 280 299 9 (POTENTIAL).
FT DOMAIN 300 306 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 307 327 10 (POTENTIAL).
FT DOMAIN 328 342 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 343 363 11 (POTENTIAL).
FT DOMAIN 364 370 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 371 391 12 (POTENTIAL).
FT DOMAIN 392 419 EXTRACELLULAR (POTENTIAL).
SQ SEQUENCE 419 AA; 44243 MW; B5BFD97DD7B9798 CRC64;

Query Match 3.7%; Score 107; DB 1; Length 419;
Best Local Similarity 20.2%; Pred. No. 1.2;
Matches 102; Conservative 73; Mismatches 179; Indels 150; Gaps 23;

QY 16 AQLLVN---LTFGLVCLAGITVPPLLLEVGVEEKFMVNL-----GIGPVLG 64
Db 14 ATVLSPFDLLASGLMDYLPV-VFPMNAL---GTTASTIQTLTYLVLMIGAGQL-- 67
QY 65 LVCVPLGASDHWGRYGRRRPFIV-----ALSGLLSLFLIPRAGWLAGLGLCPDR 118
Db 68 -----LFGPLSD-----RLGRPPVLGGGLAYVVASMGLALT----- 99
QY 119 PLELALLIGVLIDFGC-OVCFTPEALLSDLPDRPDHCRQAYSVYAFMISLGGCLGYL 177
Db 100 --SSAEVFLGLRILOACGASACLVSTFATVDIYAGREESNVYIGILGSLAMVPVAGPL 157
QY 178 LPALDWDTSALAPYLGTOECLFGLTLTFLTCVAATLLVAEEAALGPTPEAGLSAPSL 237
Db 158 LGAL-----VDMWLGRH--AIFAFLGLGMIASAAARWEPTRV----- 195
QY 238 SPHCCPCRLAFLNGLALLPRLHQLCCRPRTLRLFLVAELCSWALMTFTLYFTDFVG 297
Db 196 -----QRVAGLWSQLL-----LPVKCLNFWLYTLCYAAGMGFFVFFFS--IA 236
QY 298 EGLYQGVPAEPGTEARRHYDEGV-RMGSGLFLQCAISLVFSL-VMDRLVQRFGTRAVY 355
Db 237 PGLMNG-----RQVSQGLFSLFATVAIAMVFTAFRGMVLPKMGSPSVL 282
QY 356 LASVAAPFVAGATCLSHSVAVVTASAAALGFTFSALQILPYTLASLYHREKQVFLPKYR 415
Db 283 RMGMGC--LIAGAVLL--AITEIWALQSVLGF-----IAPMWLVGI----- 319
QY 416 GDTGGASDESLSMTS-----PLPGPKPGAPFNGHVGAGSG-----LLPPPPALCGA 463
Db 320 ---GVATASVAPNGALRGFDHVGAVTVAVYFCLGVLGSLGTLIISLLPRNTAMPVV 375
QY 464 SACDVSVRVVVGEPTAEARVPPGR 487
Db 376 VYCLTLATVVLGLSCVSRVKGSRG 399

RESULT 15
YJ94_YEAST
ID YJ94_YEAST STANDARD; PRT; 448 AA.
AC P47159;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHETICAL 49.7 KDA PROTEIN IN RP55-ZMS1 INTERGENIC REGION.
GN YJ9124C OR J2046.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID:4932;
RN [1]

RP SEQUENCE FROM N.A.
RA Rose M., Koetter P., Entian K.D.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
CC 1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement>;
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; 249624; CAA89655.1; -.
DR SGD; S0003885; YJR124C.
KW Hypothetical protein; Transmembrane; ATP-binding.
FT TRANSMEM 51 71 POTENTIAL.
FT TRANSMEM 94 114 POTENTIAL.
FT TRANSMEM 147 167 POTENTIAL.
FT TRANSMEM 185 205 POTENTIAL.
FT TRANSMEM 261 281 POTENTIAL.
FT TRANSMEM 288 308 POTENTIAL.
FT TRANSMEM 334 354 POTENTIAL.
FT TRANSMEM 387 407 POTENTIAL.
FT TRANSMEM 417 437 POTENTIAL.
FT NP_BIND 386 393 ATP (POTENTIAL).
SQ SEQUENCE 448 AA; 49663 MW; 2319CF9DD1220C14 CRC64;

Query Match 3.7%; Score 105; DB 1; Length 448;
Best Local Similarity 21.8%; Pred. No. 1.7;
Matches 94; Conservative 62; Mismatches 163; Indels 112; Gaps 25;

QY 2 VQRLWVRLRHRRKAQILLNLLTFGLVCLAGITVPPLLLEV--GVEEK---FMTMV 56
Db 16 IKLLWASVFLR-----LLSYGL-----TNQVLTFLNAINMTEDKIGLFMSLT 58
QY 57 LGIGPVLGVCVPLGASDHWGRYGRRRPFIVALSGLLSLFLIPRAGWLAGLGLCPD 116
Db 59 LA-GDV---ICSVILTYADSW---GRRVLYVGCAMML-----LSGLVFSF 98
QY 117 PRPLELALLILGVGLL---DFGQVCFPTPL-FALLSDLPDRPDHCR-QAYSVYAFMISL 170
Db 99 SENFTLLLVFAIFGVISPSSEVGP--FKSIEEAMIAHL--SPHARPEIYAIHALVGTI 154
QY 171 GCCLGYLLPALDWDTSALAPYLGTOECC--LFGLLTLFLTCVAATLLVAEEAA-----L 223
Db 155 GSALGATIGCFVDDLLKRTGLATDLOCYKLVFLVLAFFAFCKMVMILLSDATELDGHY 214
QY 224 GPTEPAEGLSAP-SLSPHCCPCRLAFLNGLALLPRLHQLCCRPRTLRLFLVAELCSW 282
Db 215 EHTDCNEETAEPDLVDNDETAPLMRQATHPE-----ERSNKLKSTVSVLMKLLV----- 263
QY 283 MALMTFTFLFYTDVFGES-----LYOGVPAEPGTEARRHYDEGVGMGSLG---LFLQC 332
Db 264 -----IFMVDLSLGSFGMTSGMWVY-----YYSKQFLMGSLALGTLFFIT 302
QY 333 AISLVFSLVMDRLVQR-FG-TRAVYLASVAA-----FPVAAGATCLSHSVAVVTASAA 384
Db 303 QLVWASSTIPSSIIARCFCGPVRTLVLQIPSGIFSIPLPMA--KNYLPLSILFLNLHFA 360
QY 385 TGFTFSALQIL 395
Db 361 TAMDMVTPRQIL 371

Search completed: February 26, 2002, 07:58:17
Job time: 102 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2002, 07:54:50 ; Search time 19.5 Seconds
(without alignments)
2160.232 Million cell updates/sec

Title: US-09-605-783a-113
Perfect score: 2861
Sequence: 1 MVORLWVSRLLRHRKAQLL.....AIYFATQVWFKSLAKYSA 553
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	347.5	12.1	515	2 T14340	sucrose-proton tra
2	337.5	11.8	523	2 T12198	sucrose transport
3	330	11.5	516	2 J02389	sucrose transport
4	328.5	11.5	594	2 G84441	probable sucrose-p
5	324.5	11.3	525	2 S28052	sucrose transport
6	323.5	11.3	512	2 S38196	sucrose transport
7	323.5	11.3	512	2 G86360	hypothetical prote
8	317.5	11.1	507	2 S48789	sucrose transport
9	311	10.9	533	2 S43142	sucrose transport
10	305.5	10.7	501	2 T14339	sucrose-proton tra
11	297.5	10.4	428	2 S48788	sucrose transport
12	294	10.3	510	2 S38657	sucrose transport
13	294	10.3	523	2 S31114	sucrose-proton sym
14	292.5	10.2	512	2 P96741	probable sucrose t
15	292	10.2	537	2 T02982	probable sucrose t
16	286	10.0	474	2 A86234	hypothetical prote
17	281.5	9.8	491	2 E96691	probable sucrose-p
18	279.5	9.8	513	2 S38197	sucrose transport
19	276.5	9.7	492	2 A84520	probable sucrose-p
20	273	9.5	503	2 S52377	probable sucrose-p
21	235.5	8.2	553	2 T38541	sucrose transport
22	191	6.7	452	2 F75217	hypothetical prote
23	142	5.0	544	2 S75636	melibiose carrier
24	139.5	4.9	454	2 A75444	hypothetical prote
25	138.5	4.8	430	2 E75217	transporter PAB217
26	133	4.6	389	2 G83413	probable MFS trans
27	128.5	4.5	422	2 G83503	probable MFS trans
28	128.5	4.5	594	2 A83096	probable permease
29	127.5	4.5	472	2 F82639	resistance protein

30 127 4.4 451 2 E81781 probable integral
31 126 4.4 451 2 A81206 sugar transporter,
32 125.5 4.4 399 2 J01479 tetracycline resis
33 125 4.4 391 2 S74688 hypothetical prote
34 125 4.4 407 2 H75515 tetracycline-efflu
35 124.5 4.4 399 1 S38656 teta protein - Pse
36 121.5 4.2 503 2 S43017 puromycin resistan
37 120.5 4.2 414 2 F83119 probable transport
38 120.5 4.2 466 2 A86033 probable permease
39 119.5 4.2 399 1 YTECR1 tetracycline resis
40 119.5 4.2 491 2 C85540 regulates beta-lac
41 119.5 4.2 491 2 S37391 signal transducer
42 119 4.2 409 2 F83493 probable MFS trans
43 119 4.2 440 2 D86025 probable transport
44 117.5 4.1 403 2 E84063 multidrug resistan
45 117.5 4.1 448 2 H83335 probable MFS trans

RESULT 1
T14340
sucrose-proton transport protein - carrot
N:Alternate names: sucrose/H+ symporter protein
C:Species: Daucus carota (carrot)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T14340
R:Shakya, R.; Sturm, A.
Plant Physiol. 118, 1473-1480, 1998
A:Title: Characterization of source- and sink-specific sucrose/H+ symporters from
A:Reference number: Z17991; MUID:99063785
A:Accession: T14340
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-515 <SHA>
A:Cross-references: EMBL:Y16768; NID:g2969883; PIDN:CAA76369.1; PID:g2969884
A:Experimental source: cultivar Nantaise; root
C:Genetics:
A:Note: SUT2
C:Superfamily: common tobacco sucrose transport protein

ALIGNMENTS

G84441

probable sucrose-proton symporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: G84441
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Mofat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: G84441
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-594 <STO>
A:CROSS-references: GB:AE002093; NID:g3461813; PIDN:AAC32907.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g02860
A:Map position: 2
C:Superfamily: common tobacco sucrose transport protein

Query Match 11.5%; Score 328.5; DB 2; Length 594;
Best Local Similarity 24.1%; Pred. No. 2.9e-17;
Matches 120; Conservative 72; Mismatches 174; Indels 131; Gaps 18;
QY 22 NLLTFGLEVCAGITY-----VPLLLEVGVEEKFMTWVLGIPVGLVGVCPVLLGS 73
DB 59 SLVTLVSCITAAVGVFGWALQSLTTPYIOTLIGISHAFSSFIWLGPIGLVGVQFVGI 118
QY 74 ASHWGRGRYRRPFTWALSGLILLSLFLIPRAGWLA--GLLCPD-----PRPLE 121
DB 119 WSDKCTSKYGRPRPFTWALSGLILLSLFLIPRAGWLA--GLLCPD-----PRPLE 175
QY 122 LALLILGVGLDFCGVCFPLEALLSDFRDPDCHCRQAYSVYAFMISLGGCLGYLLPAI 181
DB 176 AVVFIIGFWLLDIANTVQGPAPALLADL--SGPDQRTANAVFLWMAIGNILGFSAGAS 234
QY 182 ----DWTSALAPYLQTECELF-----LLTIFIT-CVAAPLLVAEEAALGPTTP 228
DB 235 GKQEW-----FPFL-TSRACCAACGNLKAFLAVVLTICTLVITYFAKEIPTTSNK 288
QY 229 AE-GLSAPLSLPHCCPCRLAFRNL-----LQSKGLEHKLNNGTANGIKYERVDTEQFGSNEHQ 253
DB 289 TRIQDSAPLDD-----LQSKGLEHKLNNGTANGIKYERVDTEQFGSNEHQ 340
QY 254 -----GALLPRLHQLCCMRPTLRLRFLVAELCSWMLMTFTLYTDFVGEGLYQGV 305
DB 341 DETYVDGPGSVLVNLLTSRLHPLPAMHSLIVMALTWLSWFPFLLPDTDMGREGVYHGP 400
QY 306 RAEPGTEARRHYDEGVMSGLFLQCAISLVFVMDRLVQRFGRVY--LASVAAPV 364
DB 401 TGD--SLHMELYDQGVREGALGULLSVLVGIFSSFIPEMCRMGARVYVWALSNTVFAC 458
QY 365 AAGATCLS-----HSVAVVTASAAALTGFTFSALQILPYTLASLY 403
DB 459 MACTAVISLMSLDDKNGIETIMRGNETTRTAIV--FALLGFLPITVSPFVTA-- 514
QY 404 HREKQVFLPKYRGDTGG 420
DB 515 -----EVTADSGS 522

RESULT 5

S28052
sucrose transport protein - spinach
N:Alternate names: sucrose carrier protein; sucrose permease
C:Species: Spinacia oleracea (spinach)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jul-2000
C:Accession: S28052
R:Riesmeier, J.W.; Willmitzer, L.; Frommer, W.B.
EMBO J. 11, 4705-4713, 1992
A:Title: Isolation and characterization of a sucrose carrier cDNA from spinach by functi

A:Reference number: S28052; MUID:93099843

A:Accession: S28052
A:Molecule type: mRNA
A:Residues: 1-525 <RIE>
C:CROSS-references: EMBL:X67125; NID:g21318; PIDN:CAA47604.1; PID:g21319
C:Superfamily: common tobacco sucrose transport protein
C:Keywords: transmembrane protein

Query Match 11.3%; Score 324.5; DB 2; Length 525;
Best Local Similarity 23.2%; Pred. No. 5.2e-17;
Matches 129; Conservative 91; Mismatches 196; Indels 141; Gaps 20;
QY 23 LLTFGLEVCAGITY-----VPLLLEVGVEEKFMTWVLGIPVGLVGVCPVLLGS 74
DB 35 LKKLGLVASAAGVFGWALQSLTTPYVOLLGIPHTWAAIWLCSPTSGMIVQPLVGY 94
QY 75 SDHWGRGRYRRPFTWALSGLILLSLFLIPRAGWLA--GLLCPD-----RPLEALLL 127
DB 95 SDRCTSRFRRRPFTAAAGALVAVAGLI---GFAADIGAASGDPGNVAKPRAIVFV 151
QY 128 GVGLLDFCGVCFPLEALLSDFR--DPDCHCRQAYSVYAFMISLGGCLGYLLPAID--- 182
DB 152 GFWLLDVANNTLQGPCRALLADMAAGSQTKTRYANAFSFFWALGNIGGYAAGSYRLYT 211
QY 183 ----WDTALAPYLQTECELFGLT--LIFLTCVAATLLVAEEAALGPTPEAGLSAPSL 237
DB 212 VPFTKTAACDVYCANLSCFPISITLITLILSVVVKERQITIDEIQEEDLKNRNN 271
QY 238 SPHCCPCRLAFRNLGALLPRLHQLCCMRPTLRLRFLVAELCSWMLMTFTLYTDFVG 297
DB 272 SSG-----ARLPF--FGQLIGALKDL----PAKMLILLVLTALNIWAFPLFTDWMG 321
QY 298 EGLYQGVPRAPETGTEARRHYDEGVMSGLFLQCAISLVFVMDRLVQRFGR--TRAVY 355
DB 322 KEVYGGT-----VGEGLYDQGVHAGALGLMINSVLGVMSLSIEGLARMVGGAKRLWG 375
QY 356 LASVAAPVAAATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYR 415
DB 376 IVNII-----AVCLAMTV-LVTKSAE-----HFRDHHIM----- 405
QY 416 GDTGSGSDESLMTSFLPGPKPGFPNGHVAGGSGLLPPPPALCGASACSVYVVG 475
DB 406 -----GSAVPPPPPA--GVKGGALAIFAVLG 429
QY 476 EPTARV-VP-----GRGICLDLAILDSFLLSQV-----AP 506
DB 430 IPLAITFTIPFALASIFSASSGSGGLSLGLVNLAIIVPQMFVSVTSGPMDAMFGGNLP 489
QY 507 SLFMGSIYQLSOSVTAY 523
DB 490 AFVGVAAVATASAVLSF 506

RESULT 6

S38196
sucrose transport protein SUC2 - Arabidopsis thaliana
N:Alternate names: sucrose-proton symporter SUC2
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Dec-2000
C:Accession: S38196; T00773
R:Sauer, N.K.
submitted to the EMBL Data Library, October 1993
A:Description: SUC1 and SUC2: two sucrose transporters from Arabidopsis thaliana.
A:Reference number: S38196
A:Accession: S38196
A:Molecule type: mRNA
A:Residues: 1-512 <SAU>
A:CROSS-references: EMBL:X75382; NID:g407091; PID:g407092
R:Vyotskaya, V.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Oji, O.; Kwan, A.; Liu, S.
rtz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federsps
submitted to the EMBL Data Library, July 1998
A:Description: Arabidopsis thaliana chromosome 1 BAC T22J18 sequence, complete seg

A:Reference number: Z14202

A:Accession: T00773

A>Status: translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-490, 'A', 492-512 <V>S>

A:CROSS-references: EMBL:AC003979; NID:g3172156; PID:g3287687; GSPDB:GN00059; ATSP:T22J1

A:Experimental source: cultivar Columbia

C:Genetics:

A:Gene: SUC2; ATSP:T22J18.12

A:Map position: 1

A:Introns: 419/3; 441/1; 455/3

C:Superfamily: common tobacco sucrose transport protein

C:Keywords: sugar transport

Query Match 11.3%; Score 323.5; DB 2; Length 512;
Best Local Similarity 24.1%; Pred. No. 6e-17;
Matches 123; Conservative 92; Mismatches 195; Indels 101; Gaps 18;

QY 14 RKAQLLVNLLTFGLVCLAAAGITYVPPLLLEVGVEKEKMTVLGIPVGLVCPVPLGS 73
DB 28 RLKLIISVSSIAAGVQFGWALQSLTTPYVQLGIPHKWASLWLCGPISGMLVQPIVG 87
QY 74 ASDHWRGRRRRPFIWALSGLILLSLFLIPRAGWLAGL-----LCPDPRPLELALLI 126
DB 88 HSDRCTSRGRRRPFIVAGAGLVAVFLI---GYAADIGHSGMDQDLDPKPKTRAIIFA 144
QY 127 LGVGLDFCGQVCFPLEALLSDFR-DPDHCRQAYSVAFMISLGGCLGY-----LL 178
DB 145 LGFWILDVANNTLQGPCRAFLADLSAGNAKKTRTANAFFSFFMAVGNVLGYAAGSYRNLY 204
QY 179 PAIDWD-TSALAPYLGTQECFLGLTLFLITCVAATLLVAEEAALGPTPEAGLSAPSL 237
DB 205 KVPFTWTESCDLYCANLTCFFLSITLL-LIVTFVSLCVVKEKPTPEADGKA---- 259
QY 238 SPHCCPCRARLAFNLGALLPRHLCCRPRTLRLFLVAELCSWMAIMTFTFLFYDFVG 297
DB 260 -----SNVPF--FGEIFGAFKEL-----KRPMMMLLIVTALNWIAPFPFLFDTDW 305
QY 298 EGLYGVPRAPGPTGTEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVORFGTRAVILA 357
DB 306 REVYGGNSDATATAASKKLYNDGVRAGALGLMLNAIVLGFMSLGVEMIGRKLK-GAKRLW 364
QY 358 SVAAPFVAAGATCLSHSVAV-----VTASA-----ALTG-----FT 388
DB 365 GIVNFIL---AICLAMTVVTKOENHRDHGGAKTGPPGNVTAGALTFLAILGIPQAIT 421
QY 389 FSALQILPYTLASLYHREK-----OVFLPKYRGDTGGASSEDLSMTSPLGP 435
DB 422 FS-----IPFALASIFSTNSGAGGSLGLVNLAIIVVPMQVVISVGGGPFDELFGGNI 476
QY 436 KPGAPFNGHVAGGSGLL-----PPPPA 459
DB 477 -----FVLGAIAAASVGLGLTVLPSPPPDA 502

RESULT 7

G86360

hypothetical protein AAC25515.1 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: G86360

R:Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maity, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: G86360

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-512 <STO>

A:CROSS-references: GB:AE005172; NID:g3287687; PIDN:AAC25515.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

C:Superfamily: common tobacco sucrose transport protein

Query Match 11.3%; Score 323.5; DB 2; Length 512;
Best Local Similarity 24.1%; Pred. No. 6e-17;
Matches 123; Conservative 92; Mismatches 195; Indels 101; Gaps 18;

QY 14 RKAQLLVNLLTFGLVCLAAAGITYVPPLLLEVGVEKEKMTVLGIPVGLVCPVPLGS 73
DB 28 RLKLIISVSSIAAGVQFGWALQSLTTPYVQLGIPHKWASLWLCGPISGMLVQPIVG 87
QY 74 ASDHWRGRRRRPFIWALSGLILLSLFLIPRAGWLAGL-----LCPDPRPLELALLI 126
DB 88 HSDRCTSRGRRRPFIVAGAGLVAVFLI---GYAADIGHSGMDQDLDPKPKTRAIIFA 144
QY 127 LGVGLDFCGQVCFPLEALLSDFR-DPDHCRQAYSVAFMISLGGCLGY-----LL 178
DB 145 LGFWILDVANNTLQGPCRAFLADLSAGNAKKTRTANAFFSFFMAVGNVLGYAAGSYRNLY 204
QY 179 PAIDWD-TSALAPYLGTQECFLGLTLFLITCVAATLLVAEEAALGPTPEAGLSAPSL 237
DB 205 KVPFTWTESCDLYCANLTCFFLSITLL-LIVTFVSLCVVKEKPTPEADGKA---- 259
QY 238 SPHCCPCRARLAFNLGALLPRHLCCRPRTLRLFLVAELCSWMAIMTFTFLFYDFVG 297
DB 260 -----SNVPF--FGEIFGAFKEL-----KRPMMMLLIVTALNWIAPFPFLFDTDW 305
QY 298 EGLYGVPRAPGPTGTEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVORFGTRAVILA 357
DB 306 REVYGGNSDATATAASKKLYNDGVRAGALGLMLNAIVLGFMSLGVEMIGRKLK-GAKRLW 364
QY 358 SVAAPFVAAGATCLSHSVAV-----VTASA-----ALTG-----FT 388
DB 365 GIVNFIL---AICLAMTVVTKOENHRDHGGAKTGPPGNVTAGALTFLAILGIPQAIT 421
QY 389 FSALQILPYTLASLYHREK-----OVFLPKYRGDTGGASSEDLSMTSPLGP 435
DB 422 FS-----IPFALASIFSTNSGAGGSLGLVNLAIIVVPMQVVISVGGGPFDELFGGNI 476
QY 436 KPGAPFNGHVAGGSGLL-----PPPPA 459
DB 477 -----FVLGAIAAASVGLGLTVLPSPPPDA 502

RESULT 8

S48789

sucrose transport protein - common tobacco

C:Species: Nicotiana tabacum (common tobacco)

C>Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 26-May-2000

C:Accession: S48789

R:Buerkle, X.Y.Z.; Frommer, W.B.

submitted to the EMBL Data Library, October 1994

A:Reference number: S48787

A:Accession: S48789

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-507 <BUE>

A:CROSS-references: EMBL:X82276; NID:g575350; PID:g575351

C:Superfamily: common tobacco sucrose transport protein

Query Match

Best Local Similarity 11.1%; Score 317.5; DB 2; Length 507;

Matches 125; Conservative 87; Mismatches 203; Indels 99; Gaps 18;

QY 4 RLWVSLRLHRRKAQLLVNLLTFGLVCLAAAGITYVPPLLLEVGVEKEKMTVLGIPVGL 63

Db 23 KLM-----KIIWASIAAGVQFGWALQSLTLTPYVQLLGIPIHKFASFVWLGGPIS 72
QY 64 GLVCVPLIGSADHWGRYGRRRPFIWALSLGILLSLFLIPRAGWLAGLLCPDP-----R 118
Db 73 GMIVQPVGVYDNCSSRFRGRRGFTAAAGAAVTIAVFLIGFAADL-GHATGDPLGKGS 131
QY 119 PLELALLIIGVGLDFCGQVCTFPLEALLSDLFDRPDHCRQAYSVAFAFMISLGGCLGYLL 178
Db 132 PRAIAVVGWFLDVANMLOQPCRALLDLGGKARMTSNAFFSFVAVGNVLYAA 191
QY 179 PAID-----WDTGALAPYGTQECFLF-GLTLIFITCVAATLLVAEEAALGTEPA 229
Db 192 GYSRLCKIFPFSKTPACDIYCANLKSCEFFIAVFLLSLTILALT--VVRENELPEKDEH 249
QY 230 EGLSAPLSLPHCCPCARLAFRNLGALLPRLHQLCCMRPTRLRLFAELCSWMLMTFT 289
Db 250 E-----IDEKAGARKSVFF--FGEIFGALKDL-----PRPMWILLVTSLNWIARFPFF 297
QY 290 LEYTFVVGGLYQGVPRABPGTEARRHYDEGVRMGSLGLFLQCAISLVFSVMDRLVQRF 349
Db 298 LYDTDMAKEVYGG--KVGDG---RLYDLGVHAGALGILLSVVLGFSLSVEFLGKKI 351
QY 350 GTRAVYLASVAAPVAAATCISHSVAVV-----TASAAITGFTFSALQI-- 394
Db 352 G-GVKRLMGILNFVL---AVCMALTVLTKMAEKSRQYDAHGTLMAPTSGVKIGALTLEA 407
QY 395 -----LPYTLASLYHREK-----QVLEPKYRGDTGGASSEDSLMTS 430
Db 408 VLGIPLAVTFSPFALASIFSSNAGSGQLSLGVNLAIWVQMLVSIAGGWDDLFGGG 467
QY 431 FLPGPKPGAPFNHGVHGGGSG-----LLPPPPA 459
Db 468 NLPG-----FIVGAVAAASGILATMLPSPPA 495

RESULT 9

S43142
sucrose transport protein - castor bean
N:Alternate names: sucrose carrier
C:Species: Ricinus communis (castor bean)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 26-May-2000
C:Accession: S43142
R:Weig, A.; Komor, E.
A:Description: A sucrose carrier from Ricinus communis.
A:Reference number: S43142
A:Accession: S43142
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-533 <WEI>
A:Cross-references: EMBL:Z31561; NID:g468561; PID:g468562
C:Superfamily: common tobacco sucrose transport protein

Query Match 10.9%; Score 311; DB 2; Length 533;
Best Local Similarity 23.3%; Pred. No. 5.5e-16;
Matches 134; Conservative 95; Mismatches 225; Indels 120; Gaps 19;

QY 17 QLLLVNLLTFGLEVCVLAAGITYVPPLLLEVGVEEKFMTVMVLGIGPVGLVCVPLIGSASD 76
Db 37 KVVWVASTAAGIOFGWALQSLTLTPYVQLLGIPIHTWAAFIWLCGPISGLMVLQPIVGYHSD 96
QY 77 HWRGYRRRRPFIWALSGLILLSLFLIPRA---GWLAG--LLCPDPRPLELALLIIGVGLL 132
Db 97 RCTSRFRRRRPFIASGAFAVAFLVIGVAADLGLSLGSDLSKPKTRAIAIFVVGFWL 156
QY 133 DEFGQVCTFPLEALLSDLF-RDPDHCROAYSVAFAFMISLGGCLGYLLPAID-----W 183
Db 157 DVANNMLOGPCRALLDLISGTSQKTRTANALFSFMAVGNVLYGAAGAYTHLYKLPFFT 216
QY 184 DTSALAPYLGTEQECFLGLLTLIFITCVAATLLVAEEAALGTEPAEGL-----SAP 235

Db 217 KTTACDVYCANLKSCEFFISIVILLSLTVALSYVKEK-----PWSPDQAVDNAEDDTASQA 272
QY 236 SLSPHCCPCARL--AFRNLGALLPRLHQLCCMRPTRLRLFAELCSWMLMTFTFLFYT 293
Db 273 SSSAQMPFFGEILGAFKNL-----KRPMMWILLVTCUNWIAWFFLLFDT 318
QY 294 DFVGEGLYQGVPRABPGTEARRHYDEGVRMGSLGLFLQCAISLVFSVMDRLVQRFQTRA 353
Db 319 DMWGREVYGG--DSSGSAEQCLKYDRGRAGALGLMNSVVLGFTSLGVEVLARGVG-GV 375
QY 354 VYLASVAAPVAAATCISHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPK 413
Db 376 KRLWGVTVNFVL---AVCLAMTV-LVTQKQEST-----RR 405
QY 414 YRGDTGASSEDSLMTSFLPGPKPGAPFNHGVHAGSGGLLPPPCALCGASACDVSRVV 473
Db 406 FATVSGGAK-----VP-----LPPP-----SGVKAGALALFAV 433
QY 474 VGEPTEARV-VP-----GRGICLDLAILDSAFLLSQVAPSLFMGSIQVL--SQS 519
Db 434 MGVPQAITYSIPALASIFSNAGSGQLSLGVNLISIVIPQIMIVSVAAGPMDALFGGN 493
QY 520 VTAYMVSAAGLGLVIAIFYATQVVDKSDLAKYSA 553
Db 494 LPAFVVGAVAALASGIFALTMLPSPQDMPMSAKA 527

RESULT 10

T14339
sucrose-proton transport protein - carrot
N:Alternate names: sucrose/H+ symporter protein
C:Species: Daucus carota (carrot)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T14339
R:Shakya, R.; Sturm, A.
A:Title: Characterization of source- and sink-specific sucrose/H+ symporters from
A:Reference number: Z17991; MUID:99063785
A:Accession: T14339
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-501 <SHA>
A:Cross-references: EMBL:Y16766; NID:g2969886; PIDN:CAA76367.1; PID:g2969887
A:Experimental source: cultivar Namtaise; leaf
C:Genetics:
A:Gene: SUT1a
C:Superfamily: common tobacco sucrose transport protein

Query Match 10.7%; Score 305.5; DB 2; Length 501;
Best Local Similarity 21.6%; Pred. No. 1.3e-15;
Matches 124; Conservative 91; Mismatches 213; Indels 145; Gaps 14;

QY 18 LLLVNLTLFGLVCVLAAGITYVPPLLLEVGVEEKFMTVMVLGIGPVGLVCVPLIGSASDH 77
Db 30 LLRVASVAGIOFGWALQSLTLTPYVQELGIPHANSSIIWLCGPLSGLLVQPIVGHMSDQ 89
QY 78 WRGYGRRRRPFIWALSGLILLSLFLIPRAGWLAGLIL--CPDPRPLELALLIIGVGLLDFC 135
Db 90 CTSKYGRRRRPFIAGTAAIILAVIIIAHSADIGLLGDDTADNKTMAIVAFVIGFWILDVA 149
QY 136 GOVCTPPEALLSDLF-RDPDHCROAYSVAFAFMISLGGCLGYLLPAIDWDTLSALAPYLG 194
Db 150 NNMTGQPCRALLDLTGNDARTRVANAYFSLFMAIGNVLYG-----ATGAYS 199
QY 195 OECEFLGLTLIFLFC-----VAATLLVAEEAAL-----GPTEPEGL 232
Db 200 YKVFPSSTSSCTINCANLKSIFYIDIPIITTTTISAAKERPRISSODGPQFSEDGT 259
QY 233 SAPLSLPHCCPCARLAFRNLGALLPRLHQLCCMRPTRLRLFAELCSWMLMTFTFLFY 292
Db 260 AQ-----SGHI-----EAFWELFEGTRLLPGSVWVILLVTCUNWIGFFPILFD 305

A:Accession: S51114
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-523 <WES>
A:Cross-references: EMBL:X83850; NID:g633171; PID:g633172
C:Superfamily: common tobacco sucrose transport protein

Query Match 10.3%; Score 294; DB 2; Length 523;
Best Local Similarity 24.3%; Pred. No. 1e-14;
Matches 125; Conservative 78; Mismatches 199; Indels 112; Gaps 20;

Qy 17 QLLVNLTLTGLEVCVLAAGTYVPPLLLEVGVEEKFMTVMVIGIPVGLVCVPLLSASD 76
Db 38 KALVASTAAGVQFGWALQSLTPYVQLLPIGHTWAPIWLCGPISGMIVQPTVGYSD 97
Qy 77 HWRGRRRRPFIWALSIGILLSLFLIPRAGWLA--GLLCPD-----PRPLEALLILGV 129
Db 98 RCTSKFGRSPFI---AVGATLVGFVSLVIGFAADIGHATGDPNGNVKPRAIIVFVGF 154
Qy 130 GLLDPCGQVCTPLEALLSLDFRDP--HCRQAYSVYAFMISLGCLGY-----LLP 179
Db 155 WILDVANNTLQGPCRALLADMAAGSAQKRYANAFVFFWALNIGGYAGSYGRLYTVFP 214
Qy 180 AIDWDTALAPYLGTQECFLGLTLFLTCVAATLLVAEEAALGPTEPAEGLSAPLS 239
Db 215 FT--HTKADTCYANLKSCEFISITLLIVTILALSVRER-----PFTLDEIQEENLN 268
Qy 240 HCCPCRLAPRNLGALLPRLHQLCCMRPTLRLFLVAELCSMMALMTFTLFYDFVGE 299
Db 269 NTGEC--ARLPF--FGQLFGALKDL---PKPMLILLVTCINWIAWFPFLFDFDWMG 321
Qy 300 LYQGVPRAPCTEARRHYDEVRMGSGLFLQCAISLVSMDRLVQRFQ--TRAVYLA 357
Db 322 VYGTG-----VGEKAYDMGVHAGLGLMINSVYVGLGMSLIGIEKLARLVGVKRLMG 375
Qy 358 SVAAPFVAAGATCLSHSVAVTASA-----ALTFG 387
Db 376 NLIL-----AVCLAMTI-LVTKSAHYRATHVHVGAIQPPPLPPGVKGNALFAVLGI 428
Qy 388 TFSALQILPYTLASLYHREK-----QVFLPKYRGDTGGASSEDLSMT 429
Db 429 PLATTFSPFALASIFSASSGGSLGVLNLAIVVQMFVSVTSGPW-----DAL-- 480
Qy 430 SFLPCPKGAPPNGHVGAGSG-----LPPPPP 458
Db 481 -FGGGLNLP--FVVGAVAAATASAILSTLPPPPP 511

RESULT 14
F96741
Probable sucrose transport protein F17M19.4 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
A:Accession: F96741
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: F96741
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-512 <STO>
A:Cross-references: GB:AE005173; NID:g6978914; PID:AAF34306.1; GSPDB:GN00141
C:Genetics:
A:Gene: F17M19.4

A:Map position: 1
C:Superfamily: common tobacco sucrose transport protein

Query Match 10.2%; Score 292.5; DB 2; Length 512;
Best Local Similarity 23.8%; Pred. No. 1.3e-14;
Matches 120; Conservative 87; Mismatches 211; Indels 87; Gaps 18

Qy 17 QLLVNLTLTGLEVCVLAAGTYVPPLLLEVGVEEKFMTVMVIGIPVGLVCVPLLSASD 76
Db 33 KIISVASIAGVQFGWALQSLTPYIQLLGIPIKWSYMWLQGPISGMIVQPIVGYSD 92
Qy 77 HWRGRRRRPFIWALSIGILLSLFLIPRAGWLA-----LCPDRPLEALLILGV 129
Db 93 RCSRFRGRRRPIAAGVALVAVSVFLI---GFAADMGHSEFGDKLENKVRTRAIIFLTGF 149
Qy 130 GLLDPCGQVCTPLEALLSLDFR--DPDHCQAYSVYAFMISLGCLGY-----LL 178
Db 150 WFLDVANNTLQGPCRFLADLAAGDAKTRVANACFFFWAVGNVLYAGSYTNLHKMF 209
Qy 179 PAIDWDTALAPYLGTQECFLGLTLFLTCVAATLLVAEEAALGPTEPAEGLSAPLS 238
Db 210 PFT--MTKACDIYCANLKTCCFFLSITLLIVTFSSLYWVKD-----QWSPQGDKEEKS 263
Qy 239 PHCCPCRLAPRNLGALLPRLHQLCCMRPTLRLFLVAELCSMMALMTFTLFYDFVE 298
Db 264 -----SUFF--FGEIFGAVR---HMRKPMVLLIVINWIAWFPFLYDIDWNGR 309
Qy 299 GLYQGVPRAPCTEARRHYDEVRMGSGLFLQCAISLVSMDRLVQRFQTRAVYLA 358
Db 310 EYVGG--NSGDERSKKLYDQGVQAGALGLMNSILLGVSGLVESIGRKM--GAKRLWG 366
Qy 359 VAAPFVAAGATCLSHSVAVTASA-----SAALTGTFFSALQIL----- 395
Db 367 CVNFILAI--LMTVLVTKSAEHHREIAGPLAGPSGKAGVFSFTVLGIPLAITYS 423
Qy 396 -PYTLASLYHREKQVFLPKYRGDTGGASSEDLSMTSLPQPKP---GAPFPN---GHVG 447
Db 424 IPFALASIFSTNSGAGQSLGVLNLAICIPQIMIVSFSSGFLDAQFGGGLNLPVVGAI 483
Qy 448 AGSGGL-----PPP--PALCGA 463
Db 484 AAVSGVLATVLPSPPPDAPAMSGA 508

RESULT 15
T02982
Probable sucrose transport protein - rice
C:Species: Oryza sativa (rice)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
A:Accession: T02982
R:Hirose, T.; Imaizumi, N.; Scofield, G.N.; Furbank, R.T.; Ohsugi, R.
Plant Cell Physiol. 38, 1389-1396, 1997
A:Title: cDNA cloning and tissue specific expression of a gene for sucrose transp
A:Reference number: Z14809; MUID:98182940
A:Accession: T02982
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-537 <HR>
A:Cross-references: EMBL:D87819; NID:g2723470; PIDN:BAA24071.1; PID:g2723471
A:Experimental source: cultivar Nipponbare, leaf
C:Genetics:
A:Gene: SUT1
C:Superfamily: common tobacco sucrose transport protein

Query Match 10.2%; Score 292; DB 2; Length 537;
Best Local Similarity 26.7%; Pred. No. 1.5e-14;
Matches 115; Conservative 72; Mismatches 167; Indels 76; Gaps 19

Qy 17 QLLVNLTLTGLEVCVLAAGTYVPPLLLEVGVEEKFMTVMVIGIPVGLVCVPLLSASD 76
Db 51 RLILSGMVGQVQGWALQSLTLPYVQTLGLSHALTSMMLCGPIAGMYVPCVGLYSD 110

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2002, 07:56:15 ; Search time 14.08 seconds
(without alignments)
883.830 Million cell updates/sec

Title: US-09-605-783a-113
Perfect score: 2861
Sequence: 1 MVQRLWVSRLLRHRKAQLL.....AIYFATQVVFDRKSLAKYSA 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues
Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2861	100.0	553	4	US-09-020-956-113
2	2861	100.0	553	4	US-09-030-607-113
3	1287	45.0	255	4	US-09-071-710-36
4	1287	45.0	255	4	US-09-525-397-36
5	330	11.5	516	1	US-08-356-340-4
6	330	11.5	516	2	US-08-786-555-4
7	324.5	11.3	525	1	US-08-356-340-2
8	324.5	11.3	525	2	US-08-786-555-2
9	243	8.5	44	4	US-09-071-710-37
10	243	8.5	44	4	US-09-525-397-37
11	146	5.1	27	4	US-09-071-710-39
12	146	5.1	27	4	US-09-525-397-39
13	135	4.7	27	4	US-09-071-710-38
14	135	4.7	27	4	US-09-525-397-38
15	109.5	3.8	455	1	US-08-035-928-2
16	106.5	3.7	436	6	5432081-10
17	106	3.7	427	4	US-09-199-737-4
18	98	3.4	713	3	US-09-335-409-11
19	97.5	3.4	3011	3	US-08-811-566-20
20	97	3.4	370	4	US-09-199-737-2
21	95.5	3.3	528	2	US-08-403-852D-21
22	95.5	3.3	528	3	US-08-510-6468-22
23	95.5	3.3	528	4	US-09-231-818-21
24	95	3.3	907	3	US-08-863-102-2
25	93.5	3.3	435	6	5268463-9
26	93.5	3.3	2995	4	US-08-444-818-138
27	93.5	3.3	3011	1	US-08-440-103-36

ALIGNMENTS

RESULT 1
US-09-020-956-113
; Sequence 113, Application US/09020956
; Patent No. 6261562
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillin, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHO
; NUMBER OF SEQUENCES: 178
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/020,956
; APPLICATION NUMBER: US/09/020,956
; FILING DATE: 09-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-09-020-956-113

Query Match 100.0%; Score 2861; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 3.9e-271;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MVQRLWVSRLLRHRKAQLLNLTFGLVGLAAGITVPPLLLLVGVVEEFMTWLVIG 60
|||||

Db 1 MVQLWVSRLLRHRKAQQLLLVLLTFTGLEYCLAAAGITVYPPLLELVGVVEKFTMTVLGIG 60
QY 61 PVGLVCVPLLGASDHWGRGRRRPFIIWALSGLIGLLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVCVPLLGASDHWGRGRRRPFIIWALSGLIGLLSLFLIPRAGWLAGLLCPDPRPL 120
QY 121 ELALLILGVGLDFCGOVCFPLEALLSDILFRDPDHCROAYSIVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLDFCGOVCFPLEALLSDILFRDPDHCROAYSIVYAFMISLGGCLGYLLPA 180
QY 181 IDWTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
Db 181 IDWTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
QY 241 CCPCRLAPRNIGALLPRLHQLCCRMPTLRLFLVAELCSNMALMTFTFLFYDFVGEGL 300
Db 241 CCPCRLAPRNIGALLPRLHQLCCRMPTLRLFLVAELCSNMALMTFTFLFYDFVGEGL 300
QY 301 YQGVPRAEPTGTEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRTGTRAVYLSVA 360
Db 301 YQGVPRAEPTGTEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRTGTRAVYLSVA 360
QY 361 APVAAAGATCLSHSVAVVTASAALTGFTFSAQLIILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APVAAAGATCLSHSVAVVTASAALTGFTFSAQLIILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKPGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFPMGSIVQLSQSVTAYMWSAAGLGLVYFATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFPMGSIVQLSQSVTAYMWSAAGLGLVYFATQ 540
QY 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 2

US-09-030-607-113
; Sequence 113, Application US/09030607
; Patent No. 6262245
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
; NUMBER OF SEQUENCES: 224
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/030,607
; FILING DATE: 25-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 553 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-09-030-607-113

Query Match 100.0%; Score 2861; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 3.9e-271; Indels 0; Gaps 0;
Matches 553; Conservative 0; Mismatches 0;

QY 1 MVQLWVSRLLRHRKAQQLLLVLLTFTGLEYCLAAAGITVYPPLLELVGVVEKFTMTVLGIG 60
Db 1 MVQLWVSRLLRHRKAQQLLLVLLTFTGLEYCLAAAGITVYPPLLELVGVVEKFTMTVLGIG 60
QY 61 PVGLVCVPLLGASDHWGRGRRRPFIIWALSGLIGLLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVCVPLLGASDHWGRGRRRPFIIWALSGLIGLLSLFLIPRAGWLAGLLCPDPRPL 120
QY 121 ELALLILGVGLDFCGOVCFPLEALLSDILFRDPDHCROAYSIVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLDFCGOVCFPLEALLSDILFRDPDHCROAYSIVYAFMISLGGCLGYLLPA 180
QY 181 IDWTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
Db 181 IDWTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
QY 241 CCPCRLAPRNIGALLPRLHQLCCRMPTLRLFLVAELCSNMALMTFTFLFYDFVGEGL 300
Db 241 CCPCRLAPRNIGALLPRLHQLCCRMPTLRLFLVAELCSNMALMTFTFLFYDFVGEGL 300
QY 301 YQGVPRAEPTGTEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRTGTRAVYLSVA 360
Db 301 YQGVPRAEPTGTEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRTGTRAVYLSVA 360
QY 361 APVAAAGATCLSHSVAVVTASAALTGFTFSAQLIILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APVAAAGATCLSHSVAVVTASAALTGFTFSAQLIILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKPGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFPMGSIVQLSQSVTAYMWSAAGLGLVYFATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFPMGSIVQLSQSVTAYMWSAAGLGLVYFATQ 540
QY 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 3

US-09-071-710-36
; Sequence 36, Application US/09071710
; Patent No. 6130043
; GENERAL INFORMATION:
; APPLICANT: BILLING-MEDEL, PATRICIA
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GORDON, JULIAN
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: HODGES, STEVEN C.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: KRATOCHVIL, JON D.
; APPLICANT: ROBERTS-RAPP, LISA
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STROUPE, STEPHEN D.
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL

TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE

NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,710
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/850,713
FILING DATE: 02-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6083.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 255 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6130043e
US-09-071-710-36

Query Match 45.0%; Score 1287; DB 4; Length 255;
Best Local Similarity 100.0%; Pred. No. 7.8e-118;
Matches 255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 299 GLYGVPAEPTGTEARRHYDEGVRMGSGLGLFLOCAISLVSLVMDRLVQRTGTRAVYLAS 358
Db 1 GLYGVPAEPTGTEARRHYDEGVRMGSGLGLFLOCAISLVSLVMDRLVQRTGTRAVYLAS 60
Qy 359 VAAFPVAAGATCLSHSVAVVTASAAALGTFTSALQILPYTLASLYHREKQVFLPKYRGDT 418
Db 61 VAAFPVAAGATCLSHSVAVVTASAAALGTFTSALQILPYTLASLYHREKQVFLPKYRGDT 120
Qy 419 GGASSEDLSMTSFLPGPKGAPFNGHVAGGSGLLPPPPALCGASACDVSRRVVVGEP 478
Db 121 GGASSEDLSMTSFLPGPKGAPFNGHVAGGSGLLPPPPALCGASACDVSRRVVVGEP 180
Qy 479 EARVVPGRGICLDLAIDLDAISAFLLSQVAPSLFMGSIQVLSQSVTAYMYSAAGLGLVAYFA 538
Db 181 EARVVPGRGICLDLAIDLDAISAFLLSQVAPSLFMGSIQVLSQSVTAYMYSAAGLGLVAYFA 240
Qy 539 TQVVFDSKDLAKYSA 553
Db 241 TQVVFDSKDLAKYSA 255

RESULT 4

US-09-525-397-36
Sequence 36, Application US/09525397
Patent No. 6252047
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.

APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLASS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/525,397
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/071,710
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6083.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 255 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6252047e
US-09-525-397-36

Query Match 45.0%; Score 1287; DB 4; Length 255;
Best Local Similarity 100.0%; Pred. No. 7.8e-118;
Matches 255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 299 GLYGVPAEPTGTEARRHYDEGVRMGSGLGLFLOCAISLVSLVMDRLVQRTGTRAVYLAS 358
Db 1 GLYGVPAEPTGTEARRHYDEGVRMGSGLGLFLOCAISLVSLVMDRLVQRTGTRAVYLAS 60
Qy 359 VAAFPVAAGATCLSHSVAVVTASAAALGTFTSALQILPYTLASLYHREKQVFLPKYRGDT 418
Db 61 VAAFPVAAGATCLSHSVAVVTASAAALGTFTSALQILPYTLASLYHREKQVFLPKYRGDT 120
Qy 419 GGASSEDLSMTSFLPGPKGAPFNGHVAGGSGLLPPPPALCGASACDVSRRVVVGEP 478
Db 121 GGASSEDLSMTSFLPGPKGAPFNGHVAGGSGLLPPPPALCGASACDVSRRVVVGEP 180
Qy 479 EARVVPGRGICLDLAIDLDAISAFLLSQVAPSLFMGSIQVLSQSVTAYMYSAAGLGLVAYFA 538
Db 181 EARVVPGRGICLDLAIDLDAISAFLLSQVAPSLFMGSIQVLSQSVTAYMYSAAGLGLVAYFA 240
Qy 539 TQVVFDSKDLAKYSA 553
Db 241 TQVVFDSKDLAKYSA 255

```
RESULT 5
US-08-356-340-4
; Sequence 4, Application US/08356340
; Patent No. 5608146
; GENERAL INFORMATION:
; APPLICANT: FROMMER, Wolf-Bernd
; APPLICANT: RIESMEIER, Jorg
; TITLE OF INVENTION: DNA SEQUENCES WITH OLIGOSACCHARIDE
; TITLE OF INVENTION: TRANSPORTER, PLASMIDS, BACTERIA AND PLANTS CONTAINING A
; TITLE OF INVENTION: TRANSPORTER AS WELL AS A PROCESS FOR THE PREPARATION AND
; TITLE OF INVENTION: TRANSFORMATION OF YEAST STRAINS FOR THE IDENTIFI...
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
; STREET: 1180 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10036-8403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,340
; FILING DATE: 21-DEC-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/EP93/01604
; FILING DATE: 22-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P42 20 759.2
; FILING DATE: 24-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Meilman, Edward A.
; REGISTRATION NUMBER: 24,735
; REFERENCE/DOCKET NUMBER: P/951-106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 382-0700
; TELEFAX: (212) 382-0888
; TELEX: 236925
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 516 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-356-340-4

Query Match 11.5%; Score 330; DB 1; Length 516;
Best Local Similarity 24.7%; Pred. No. 6.9e-24;
Matches 132; Conservative 93; Mismatches 197; Indels 112; Gaps 21;

QY 4 RLWVSRLLRHRKAQLLLVNLFTFGLVCLAGITYVPPLLLLEVGVEEKFMTVMVLGIPVL 63
DB 29 KLV-----KIIIVASTAAGVQFGWALQSLTTPVQLLGIPIHKFASFIVLWCGPIS 78

QY 64 GLVCVPLLGASDHWGRYGRRRPFIWALSGLLSLFLIPRAGWLAGL-----LCPD 116
DB 79 GMIVQPVVGYYSNDCSSRFGRRRPPIAAGALVMIAVFLI---GFAADLGHASGDTLGK 135

QY 117 PRPLEALLILGVLLDPCGQVCFPTLEALLSDLFDPD-HCRQAYSVYAFMISLGGCLG 175
DB 136 FKPRAIADVGVFWILDVANNMLOGPCRALLADLSGGSGRMTANAFESFFMAVGNILG 195

QY 176 YLLPAIDW-----DTSALAPYLGTQECLEF-GLLTILFLTCVAATLLVAEE-AALGP 225
DB 196 YAAGSYHLKVPFSPKTKACDMYCANLKSCFFIAIFLLLSLTTLTILVRENELPERKE 255

QY 226 TEPAGLSAPLSPHCCPCPCARLAFNLGALLPRHLQCCMRPTRLRLFLVAELCSWAL 285
DB 256 QEIDEKLAGAG-----KSKVPF--FGEIFGALKEL-----PRPMWILLVTLCLNWI 301

QY 286 MTFTLFYTDVFGEGLYQGVPRAEPCGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRL 345

US-08-786-555-4
; Sequence 4, Application US/08786555B
; Patent No. 5981181
; GENERAL INFORMATION:
; APPLICANT: FROMMER, Wolf-Bernd
; APPLICANT: RIESMEIER, Jorg
; TITLE OF INVENTION: DNA SEQUENCES WITH OLIGOSACCHARIDE TRANSPORTER,
; TITLE OF INVENTION: PLASMIDS, BACTERIA AND PLANTS CONTAINING A TRANSPORTER
; TITLE OF INVENTION: AS WELL AS A PROCESS FOR PREPARATION AND TRANSFORMATION
; TITLE OF INVENTION: OF YEAST STRAINS FOR THE IDENTIFICATION
; FILE REFERENCE: Frommer
; CURRENT APPLICATION NUMBER: US/08/786,555B
; CURRENT FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: 08/356,340
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: PCT/EP93/01604
; EARLIER FILING DATE: 1993-06-22
; EARLIER APPLICATION NUMBER: DE P4220759.2
; EARLIER FILING DATE: 1992-06-24
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 516
; TYPE: PPT
; ORGANISM: Spinacia oleracea
US-08-786-555-4

Query Match 11.5%; Score 330; DB 2; Length 516;
Best Local Similarity 24.7%; Pred. No. 6.9e-24;
Matches 132; Conservative 93; Mismatches 197; Indels 112; Gaps 21;

QY 4 RLWVSRLLRHRKAQLLLVNLFTFGLVCLAGITYVPPLLLLEVGVEEKFMTVMVLGIPVL 63
DB 29 KLV-----KIIIVASTAAGVQFGWALQSLTTPVQLLGIPIHKFASFIVLWCGPIS 78

QY 64 GLVCVPLLGASDHWGRYGRRRPFIWALSGLLSLFLIPRAGWLAGL-----LCPD 116
DB 79 GMIVQPVVGYYSNDCSSRFGRRRPPIAAGALVMIAVFLI---GFAADLGHASGDTLGK 135

QY 117 PRPLEALLILGVLLDPCGQVCFPTLEALLSDLFDPD-HCRQAYSVYAFMISLGGCLG 175
DB 136 FKPRAIADVGVFWILDVANNMLOGPCRALLADLSGGSGRMTANAFESFFMAVGNILG 195

QY 176 YLLPAIDW-----DTSALAPYLGTQECLEF-GLLTILFLTCVAATLLVAEE-AALGP 225
DB 196 YAAGSYHLKVPFSPKTKACDMYCANLKSCFFIAIFLLLSLTTLTILVRENELPERKE 255

QY 226 TEPAGLSAPLSPHCCPCPCARLAFNLGALLPRHLQCCMRPTRLRLFLVAELCSWAL 285
DB 256 QEIDEKLAGAG-----KSKVPF--FGEIFGALKEL-----PRPMWILLVTLCLNWI 301

QY 286 MTFTLFYTDVFGEGLYQGVPRAEPCGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRL 345
```

```

Db   302 FPFELYTDWMAKEVFGG---QVGD---ARLYDLGVAGAMGLLLQSIVLGFMSLGVEFL 355
      :|::||:::||::||::||::||::||::||::||::||::||::||::||::||::||
Qy   346 VOREGTAVYLASVAAPFAAGATCLSHSVAVV-----TASAALTGET----- 388
      ::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   356 GKKIG-GAKRWGLNLFVL---ALCLAWTILVTWKAEKSRQHDPAGTLMGPTPGVKIGAL 411
      :|::||::||::||::||::||::||::||::||::||::||::||::||
Qy   389 --FSALQI-----LPYTLASHYHREK-----QVFELPKRYRGGDTGGSSEDS 426
      :|::||::||::||::||::||::||::||::||::||::||::||::||
Db   412 LLFPAALGPLAATESIFFALASIFSSNRSGQLSLGLVLNLAIIVVPQMVLVSLVGPPDDL 471
      :|::||::||::||::||::||::||::||::||::||::||::||::||
Qy   427 LMTSELFGRPGAFPNGHVAGGSG-----LLPPPALCGACASDVSVRVVVG 475
      :|::||::||::||::||::||::||::||::||::||::||::||::||
Db   472 FGGGNLFC-----FVYGAVAARAASAYLAULTMLPSPA-----DAKPAVAMG 512
      :|::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 7
US-08-356-340-2
; Sequence 2, Application US/08356340
; Patent No. 5608146
; GENERAL INFORMATION:
; APPLICANT: FROMMER, Wolf-Bernd
; APPLICANT: RIESMEIER, Jorg
; TITLE OF INVENTION: DNA SEQUENCES WITH OLIGOSACCHARIDE
; TITLE OF INVENTION: TRANSPORTER, PLASMIDS, BACTERIA AND PLANTS CONTAINING A
; TITLE OF INVENTION: TRANSPORTER AS WELL AS A PROCESS FOR THE PREPARATION AND
; TITLE OF INVENTION: TRANSFORMATION OF YEAST STRAINS FOR THE IDENTIFI...
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
; STREET: 1180 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10036-8403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,340
; FILING DATE: 21-DEC-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/EP93/01604
; FILING DATE: 22-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P42 20 759.2
; FILING DATE: 24-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Meilman, Edward A.
; REGISTRATION NUMBER: 24,735
; REFERENCE/DOCKET NUMBER: P/951-106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 382-0700
; TELEFAX: (212) 382-0888
; TELEX: 236925
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 525 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-356-340-2

```

```

Query Match      11.3%  Score 324.5;  DB 1;  Length 525;
Best Local Similarity 23.2%;  Pred. No. 2.4e-23;
Matches 129;  Conservative 91;  Mismatches 196;  Indels 141;  Gaps 20;

QY  23  LLTGLEVCLAAQTY-----VPLLLEVGVEERFMTWVLGIGPVLGVCVPLGSA 74
      |  ||  :|||:  :  :  :||:  :  :  :  ||:  :||:  :||:

```

D	b	35	LKKGLGVASVAGVQGNALQSLLTPYVQLLGPHTWAAYIWLCPISGMIVQPLVGYY	94
Q	y	75	SDHWGRGYGRRRPFIWALSGLILLSLFIPRAGWLA--GLLCPPDP-----RPLELALLIL	127
D	b	95	SDBCTSGREGRPFIAAALVAVAVGLI---GFAADIGAASGDPTGNVAKPRAITAVEVV	151
Q	y	128	GVGLDFCGQVCFTPLEALLSDLFR-DPDHCRQAYSVAFMISLGCGCLGYLLPAID----	182
D	b	152	GFWILDVANNTLQGPCRRALLADMMAAGSQTKTRYANAFESFFMALNIGGYAAGSYSRLYT	211
Q	y	183	---WDTSALAPYLGTQECLFGLLT-LIFLTCVAATLLVAEEAALGTEPAEGLSAPSLS	237
D	b	212	VFPFTKTACDVCANLKSCFFISTLIIVLITLALSVMVKERQITIIDEIOBEDLKNRNN	271
Q	y	238	SPHCPCRARLARNLGALLPRHLHOLCCMRPTLRRLFAELCSWMALMTFTLFYTDVFG	297
D	b	272	SSGC----ARLPF--FGOLIGALKDL----PKPMLTILLVLVTALNINIAWFEPFLFFDTDMWG	321
Q	y	298	EGLYOGVPRAEPGTEARRHYDEGVWGSIGLFLQCAISLVFSLVMDRLVOREG--TRAVY	355
D	b	332	KEVYGGT-----VGEGLKYDOGVHAGAUGLMINSVLGVMSLSLEGLARMVGGAKRLWG	375
Q	y	356	LASVAAFPVAAGATCLSHSVAVVTASAALTGTFESAOLILPYTLASLYHREKQVLPKYR	415
D	b	376	IVNIIL-----AVCLAMTV-LVTKSAE-----HFDRSHHIM-----	405
Q	y	416	GDTGGASSEDSMTSEFLPGPKGAPFNHGVGAGSGGLLPPPPALCGSACDVSVRVVVG	475
D	b	406	-----GSAVPPPPPPA--GVKGGA LAIFA VLG	429
Q	y	476	EPTEARV-VP-----GRGICLDLATILDS AFLLSQV-----AP	506
D	b	430	IPLATIFSIPUALASIPSASSGGOSLGLVNLAIVPQMFEVSYTSGPWDAMFGGNLP	489
Q	y	507	SLEMGSIVOLSQSVTAY	523
D	b	490	AFVVGAAVATAASAVLSF	506

RESULT 8

```

US-08-786-555-2
; Sequence 2, Application US/08786555B
; Patent No. 5981181
; GENERAL INFORMATION:
; APPLICANT: FROMMER, Wolf-Bernd
; APPLICANT: RIESMEIER, Jorg
; TITLE OF INVENTION: DNA SEQUENCES WITH OLIGOSACCHARIDE TRANSPORTER,
; TITLE OF INVENTION: PLASMIDS, BACTERIA AND PLANTS CONTAINING A TRANSPORTER
; TITLE OF INVENTION: AS WELL AS A PROCESS FOR PREPARATION AND TRANSFORMATION
; TITLE OF INVENTION: OF YEAST STRAINS FOR THE IDENTIFICATION
; FILE REFERENCE: Frommer
; CURRENT APPLICATION NUMBER: US/08/786.555B
; CURRENT FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: 08/356.340
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: PCI/EP93/01604
; EARLIER FILING DATE: 1993-06-22
; EARLIER APPLICATION NUMBER: DE P4220759.2
; EARLIER FILING DATE: 1992-06-24
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 525
; TYPE: PRt
; ORGANISM: Spinacia oleracea
; US-08-786-555-2
    
```

Query Match 11.3%; Score 324.5; DB 2; Length 525;
Best Local Similarity 23.2%; Pred. No. 2.4e-23;
Matches 129; Conservative 91; Mismatches 196; Indels 141; Gaps 20

Qv 23 LLTFGLVCLAAQITY-----VPPILLLEVGVSEKFTMWVLGPGVLGLVCPVLLGSA 74

Db	35	LKXGLVSAVAAQVQFGWALQJLSLTPPYVOLJGIPHTWAAIYWLQCGIPSGMTVPLQVGY	94
Qy	75	SDHWGRYGRRRRPFIWALSGILLSLFLPRAGWLA--GLLCPDP-----RPLELALLIL	127
Db	95	SDRCTSEFGRRRPFIAAGAALVAVAGCLI--GFAADIGAASGDPGTGNVAKPRAVAFV	151
Qy	128	GVGLLDFCGQVCTPTEALLSDLFR--DPHCRQOAVSVYAFMISLGCGLGYLLPAID----	182
Db	152	GFWLIDVANNITLOGPCRALIADMAAGSQTTRYANAFFSFFMALGNIGGYAAGSYSRLYT	211
Qy	183	----WOTSALAPYLGTOEECLFGLLT-LIFLTCVAATILVAEEAALGPTEPAEGLSAPSL	237
Db	212	VFETTKAACDVTCANKSCFFISITLLIVLTILALSVVYKEROITIDEIOEEDLKNRN	271
Qy	238	SPHCCPCRAAFAPRNLGALLPRHLQCCRMPTRLRFLVAELCSMMALMTFTFLTDFVG	297
Db	272	SSGC-----ARLPE-FQLIGALKDL---PKPMLILLVLTALNWTIANWFPFLFDTDMG	321
Qy	298	EGLYQGVPRAPGCTEARRHYDEQVMSGLFLQCAISLVFSLVMDRLVORFG--TRAVY	355
Db	322	KEYVGGF-----VGEKLYDQGVHAGALGLMTNSVVLGVMSLSTIEGLARMVGGAKRLWG	375
Qy	356	LASVAAPVAAGATCLSHSVAVVTASAALTGFTFSALQILPTLASLYHREKQVFLPKYR	415
Db	376	IVNLIIL-----AVCLAMTV-LVTKSAE-----HFRDSHIM-----	405
Qy	416	GDTGCGASSEDLSMTSLPQPKGPPNGHVAGSGGLLPPPPALCCGASACDVSVRVVVG	475
Db	406	-----GSAVPPPPPA--GVKGGALAIFAVLG	429
Qy	476	EPTEARV-VP-----GRGICLDLAILDASFLLSQV-----AP	506
Db	430	IPLAITEIFLALASTFSSASSGGSLGLVNLAIYVQMEVSVTSGPWDAMFGGNLP	489
Qy	507	SLFMGSTVQLSQSVTAY	523
Db	490	AFVVGAVAATASAVLSF	506
RESULT 9			
US-09-071-710-37			
; Sequence 37, Application US/09071710			
; Patent No. 6130043			
; GENERAL INFORMATION:			
; APPLICANT: BILLING-MEDEL, PATRICIA			
; APPLICANT: COHEN, MAURICE			
; APPLICANT: COLPITTS, TRACEY L.			
; APPLICANT: FRIEDMAN, PAULA N.			
; APPLICANT: GORDON, JULIAN			
; APPLICANT: GRANADOS, EDWARD N.			
; APPLICANT: HODGES, STEVEN C.			
; APPLICANT: KLASS, MICHAEL R.			
; APPLICANT: KRATOCHVIL, JON D.			
; APPLICANT: ROBERTS-RAPP, LISA			
; APPLICANT: RUSSELL, JOHN C.			
; APPLICANT: STROUPE, STEPHEN D.			
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL			
; FOR DETECTING DISEASES OF THE PROSTATE			
; NUMBER OF SEQUENCES: 41			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Abbott Laboratories			
; STREET: 100 Abbott Park Road			
; CITY: Abbott Park			
; STATE: IL			
; COUNTRY: USA			
; ZIP: 60064-3500			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Diskette			
; COMPUTER: IBM Compatible			
; OPERATING SYSTEM: DOS			
; SOFTWARE: FastSeq for Windows Version 2.0			
; CURRENT APPLICATION DATA:			

NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6083.US.PI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:

INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 44 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6252047e
US-09-525-397-37

Query Match 8.5%; Score 243; DB 4; Length 44;
Best Local Similarity 100.0%; Pred. No. 6.4e-17;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 403 YHREKQVFLPKYRGDTGGASDLSMTSLPGPKGAPFNGHV 446
|||||
DB 1 YHREKQVFLPKYRGDTGGASDLSMTSLPGPKGAPFNGHV 44

RESULT 11
US-09-071-710-39
; Sequence 39, Application US/09071710
; Patent No. 6130043
; GENERAL INFORMATION:
; APPLICANT: BILLING-MEDEL, PATRICIA
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GORDON, JULIAN
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: HODGES, STEVEN C.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: KRATOCHVIL, JON D.
; APPLICANT: ROBERTS-RAPP, LISA
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STROUPE, STEPHEN D.
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
; TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,710
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/850,713
; FILING DATE: 02-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6083.US.PI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:

INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6130043e
US-09-071-710-39

Query Match 5.1%; Score 146; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 9.5e-08;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 GLYQGVPRAEPCGTEARRHYDEGVRMGS 325
|||||
DB 1 GLYQGVPRAEPCGTEARRHYDEGVRMGS 27

RESULT 12
US-09-525-397-39
; Sequence 39, Application US/09525397
; Patent No. 6252047
; GENERAL INFORMATION:
; APPLICANT: BILLING-MEDEL, PATRICIA
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GORDON, JULIAN
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: HODGES, STEVEN C.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: KRATOCHVIL, JON D.
; APPLICANT: ROBERTS-RAPP, LISA
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STROUPE, STEPHEN D.
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
; TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/525,397
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/071,710
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6083.US.PI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6252047e

US-09-525-397-39

Query Match 5.1%; Score 146; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 9.5e-08;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 GLYGVPRAPGTEARRHHYDEGVRMG 325
|||||
Db 1 GLYGVPRAPGTEARRHHYDEGVRMG 27

RESULT 13

US-09-071-710-38
; Sequence 38, Application US/09071710
; Patent No. 6130043
; GENERAL INFORMATION:
; APPLICANT: BILLING-MEDEL, PATRICIA
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GORDON, JULIAN
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: HODGES, STEVEN C.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: KRATOCHVIL, JON D.
; APPLICANT: ROBERTS-RAPP, LISA
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STROUPE, STEPHEN D.
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: 08/850,713
; APPLICATION NUMBER: 08/850,713
; FILING DATE: 02-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6083.US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6130043e
US-09-071-710-38

Query Match 4.7%; Score 135; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 471 RVVVGTEARVVPGRGICLDLALDS 497
|||||
Db 1 RVVVGTEARVVPGRGICLDLALDS 27

RESULT 14

US-09-525-397-38
; Sequence 38, Application US/09525397
; Patent No. 6252047
; GENERAL INFORMATION:
; APPLICANT: BILLING-MEDEL, PATRICIA
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GORDON, JULIAN
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: HODGES, STEVEN C.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: KRATOCHVIL, JON D.
; APPLICANT: ROBERTS-RAPP, LISA
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STROUPE, STEPHEN D.
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: 09/071,710
; APPLICATION NUMBER: 09/071,710
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6083.US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6252047e
US-09-525-397-38

Query Match 4.7%; Score 135; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 471 RVVVGTEARVVPGRGICLDLALDS 497
|||||
Db 1 RVVVGTEARVVPGRGICLDLALDS 27

RESULT 15

US-08-035-928-2

; Sequence 2, Application US/08035928
; Patent No. 5538844
; GENERAL INFORMATION:
; APPLICANT: Duyao, Mabel P.
; APPLICANT: MacDonald, Marcy E.
; APPLICANT: Gusella, James F.
; TITLE OF INVENTION: A No. 5538844el Transport Protein Gene from
; TITLE OF INVENTION: the Huntington's Disease Region
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Avenue N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/035,928
; FILING DATE: 19930323
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 466-0800
; TELEFAX: (202) 833-8716
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 455 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-035-928-2

Query Match 3.8%; Score 109.5; DB 1; Length 455;
Best Local Similarity 22.9%; Pred. No. 0.02;
Matches 107; Conservative 55; Mismatches 162; Indels 143; Gaps 22;

QY	19	LLVNLFFGLEVLAAGITVV-----PPL-----LLEVGVEEKMTVLG--	58
Db	32	LLLDLFTLLPLPGLLESHGRADPLYGSGOGVDWFATAIGMPVEKRYNSVLFGL	91
QY	59	IG---PVLGLVCVPLGSSADHNRGRRRPFITWALSGLLSLFLIPRAGWLAGLLCP	115
Db	92	IGSAFVQLFCAPLTGATSDCL-----GRRPVMLCLMGVATSYAVWATSRSAFLA-	145
QY	116	DRPLEALILGLVGLDFCGQVCFPLEALLSDLFDRDPDHCQAYSVAFMISLGGCLG	175
Db	146	-----SRLLGGISK-----GNVSL--TAIVADL--GSLARSGQMAVIGVAFSLGFTLG	191
QY	176	YLLPA---IDWTSALAPYLGTOBECLFGLTLTIFTCVAATLIVAEAAALGPTPAEGL	232
Db	192	PMIGASLPLE-----MAPWFA---LLFAASDLLFICFLPETLPLEK-----	230
QY	233	SAPSLSPHCPCRARLARFNLGALLPRHLQCCRM-----RTLRRLFAE	278
Db	231	RAPSIA-----LGFRDAADLLSPALL--RFSAVARGQDPPSGDRLSLRRL--G	276
QY	279	LCSWMAIMTWT-----LFYTDVCGEGLYQG-----VPAEPCGTEA	313
Db	277	LVYFLYLFSGLEYTLSTFTHORFQSSLOQGMFFLIGLTWATOGAYARRIHGGEV	336
QY	314	RR-----HYDEGVMGSLGLFLOCAISLVSFLVMDRLVQRF-----GTR	352
Db	337	AAVKRALLLVPAFLLIWGRSLPVLGLGLLYSFAAAVVVPCLLSSVVGCGSPGQGTV	396
QY	353	AVYLASVAAPFVAAGATCLSHSVAVVTASALTGTFPSALQIIPYL	399
Db	397	MGLTSLRSLGARAAGPLVAASVYWLGAQAQCT--TWSGLFLLPFFL	441

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2002, 07:52:50 ; Search time 25.01 Seconds
(without alignments)
1637.846 Million cell updates/sec

Title: US-09-605-783a-113

Perfect score: 2861

Sequence: 1 MVQRLWVSRLLRHRKAQLL.....AIYFATQWFDKSLAKYSA 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_1101.*
- 1: /SID88/gcgdata/geneseq/geneseqp/AA1980.DAT.*
 - 2: /SID88/gcgdata/geneseq/geneseqp/AA1981.DAT.*
 - 3: /SID88/gcgdata/geneseq/geneseqp/AA1982.DAT.*
 - 4: /SID88/gcgdata/geneseq/geneseqp/AA1983.DAT.*
 - 5: /SID88/gcgdata/geneseq/geneseqp/AA1984.DAT.*
 - 6: /SID88/gcgdata/geneseq/geneseqp/AA1985.DAT.*
 - 7: /SID88/gcgdata/geneseq/geneseqp/AA1986.DAT.*
 - 8: /SID88/gcgdata/geneseq/geneseqp/AA1987.DAT.*
 - 9: /SID88/gcgdata/geneseq/geneseqp/AA1988.DAT.*
 - 10: /SID88/gcgdata/geneseq/geneseqp/AA1989.DAT.*
 - 11: /SID88/gcgdata/geneseq/geneseqp/AA1990.DAT.*
 - 12: /SID88/gcgdata/geneseq/geneseqp/AA1991.DAT.*
 - 13: /SID88/gcgdata/geneseq/geneseqp/AA1992.DAT.*
 - 14: /SID88/gcgdata/geneseq/geneseqp/AA1993.DAT.*
 - 15: /SID88/gcgdata/geneseq/geneseqp/AA1994.DAT.*
 - 16: /SID88/gcgdata/geneseq/geneseqp/AA1995.DAT.*
 - 17: /SID88/gcgdata/geneseq/geneseqp/AA1996.DAT.*
 - 18: /SID88/gcgdata/geneseq/geneseqp/AA1997.DAT.*
 - 19: /SID88/gcgdata/geneseq/geneseqp/AA1998.DAT.*
 - 20: /SID88/gcgdata/geneseq/geneseqp/AA1999.DAT.*
 - 21: /SID88/gcgdata/geneseq/geneseqp/AA2000.DAT.*
 - 22: /SID88/gcgdata/geneseq/geneseqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	ID	Description
1	2861	100.0	553	19	AAW71869	Amino acid encoded
2	2861	100.0	553	19	AAW69385	Prostate tumour sp
3	2861	100.0	553	21	AAW28527	Protein encoded by
4	2861	100.0	553	21	AAW82002	Human immunogenic
5	2861	100.0	553	22	AAW04961	Human prostate tum
6	2861	100.0	553	22	AAW01117	Human prostate-spe
7	2861	100.0	553	22	AAW99002	Human prostate-spe
8	2861	100.0	553	22	AAW62150	Human P501S invent
9	2861	100.0	553	22	AAW74800	Prostate tumour an
10	2861	100.0	1079	22	AAW74830	Prostate tumour an
11	2596	90.7	595	22	AAW01318	Alpha prepro-P501S

12	1417.5	49.5	371	22	AAW01230	P553S splice varia
13	1416	49.5	371	22	AAE01362	Human gene 11 enco
14	1403.5	49.1	400	22	AAW01262	Ra12-P501S-E2 cons
15	1287	45.0	255	20	AAW85068	Protein encoded by
16	1287	45.0	255	21	AAW29268	Human prostate-rel
17	1287	45.0	255	22	AAW04205	Prostate-specific
18	1150	40.2	231	21	AAW54369	Amino acid sequenc
19	1120	39.1	252	22	AAE01423	Human secreted pro
20	457.5	16.0	748	22	AAW04227	Human polypeptide
21	426	14.9	123	22	AAW01228	P553S splice varia
22	357	12.5	494	20	AAW41119	Soybean sucrose tr
23	337.5	11.8	523	20	AAW41129	Vicia faba sucrose
24	326.5	11.4	667	20	AAW41118	Rice sucrose trans
25	325	11.4	563	20	AAW41125	Wheat sucrose tran
26	323.5	11.3	507	21	AAW30206	Arabidopsis thalia
27	323.5	11.3	512	21	AAW30205	Arabidopsis thalia
28	323.5	11.3	539	21	AAW30204	Arabidopsis thalia
29	313	10.9	497	20	AAW41116	Corn sucrose trans
30	312	10.9	58	22	AAW01188	Human prostate-spe
31	312	10.9	58	22	AAW99073	Human prostate-spe
32	311	10.9	533	20	AAW41128	Ricinus communis s
33	305.5	10.7	501	20	AAW41126	Daucus carota suc
34	303.5	10.6	519	20	AAW41114	Corn sucrose trans
35	303.5	10.6	522	20	AAW41123	Wheat sucrose tran
36	300	10.5	522	20	AAW41124	Wheat sucrose tran
37	292	10.2	537	20	AAW41127	Oryza sativa sucro
38	291	10.2	619	21	AAW40554	Human OREF (GF318
39	289	10.1	150	22	AAW01229	P553S splice varia
40	267	9.3	267	22	AAW42013	Human polypeptide
41	243	8.5	44	20	AAW85069	PS108 protein deri
42	243	8.5	44	21	AAW29269	Human prostate-rel
43	243	8.5	44	22	AAW04206	Prostate-specific
44	235.5	8.2	322	20	AAW41120	Soybean sucrose tr
45	197.5	6.9	400	20	AAW41117	Rice sucrose trans

ALIGNMENTS

RESULT	1
AAW71869	
ID	AAW71869 standard; Protein; 553 AA.
AC	AAW71869;
XX	
DT	06-JAN-1999 (first entry)
XX	
DE	Amino acid encoded by prostate tumour clone L1-12.
XX	
KW	Prostate; cancer; tumour; vaccine; immunogen; clone.
XX	
OS	Homo sapiens.
XX	
PN	W09837093-A2.
XX	
PD	27-AUG-1998.
XX	
PF	25-FEB-1998; 98WO-US03492.
XX	
PR	09-FEB-1998; 98US-0020956.
XX	
PR	25-FEB-1997; 97US-0806099.
XX	
PR	01-AUG-1997; 97US-0904804.
XX	
PA	(CORI-) CORIXA CORP.
XX	
PI	Dillon DC, Xu J;
XX	
DR	WPI; 1998-609886/51.
XX	
DR	N-PSDB; AAW61201.
XX	
PT	Polypeptides comprising immunogenic portions of prostate proteins
XX	
PT	used in a vaccine for the treatment of prostate cancer

```
PS Example 1; Page 82-84; 130pp; English.
XX
CC The present sequence is an immunogenic portion of a prostate tumour
CC protein. The immunogen, or the DNA encoding it, can be used as a
CC vaccine for the treatment of prostate cancer. The immunogen was
CC isolated from a prostate tumour cDNA library obtained by subtracting
CC a prostate tumour cDNA expression library with a normal tissue cDNA
CC library.
XX
SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 19; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.2e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLRRKKAQOLLVNLITFGLEVCLAAGITVYPPLLLEVGVEEKFTMWLGIG 60
DB 1 mvqrlwvsrllrrhrkaqllvnlitfglevclaagityvpplllevgveekfntmvlgi 60

QY 61 PVLGVCVPLLGASDHWGRYGRRRPFIWALSIGLILSLFLIPRAGWLAGLPCDPRPL 120
DB 61 pvlglvcvplllgssadhwgrgyrrrrpfiwalsigllslflipragwlaglpcdprpl 120

QY 61 pvlglvcvpllgssadhwgrgyrrrrpfiwalsigllslflipragwlaglpcdprpl 120
DB 61 pvlglvcvpllgssadhwgrgyrrrrpfiwalsigllslflipragwlaglpcdprpl 120

QY 121 ELALLILGVGLDFCGOVCFPLEALLSDFRDPDHCROQAVSVYAFMISLGGCLGYLLPA 180
DB 121 elallilgvglldfcgvcftpleallsdfrdpdchrqaysvyafmisl9gclgyllpa 180

QY 121 elallilgvglldfcgvcftpleallsdfrdpdchrqaysvyafmisl9gclgyllpa 180
DB 121 elallilgvglldfcgvcftpleallsdfrdpdchrqaysvyafmisl9gclgyllpa 180

QY 181 IDWTSALAPYLGTOECFLGLTLTFLTCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240
DB 181 idwtsalapylgtqeecflglltfltcvaatl1lvaeeaa1gptepaeglsapslsph 240

QY 241 CCPCRLARFNLGALLPRHLQCCMRPTRLRLRFLVAELCSWALMTFTLYTDFVGEGL 300
DB 241 ccpcrlarfnlgallprhlqccmrptrlrrlrfvaelcswalmftftlytdfvgegl 300

QY 301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
DB 301 yqgvpraeptgearrhydegvrmgs1glflqcaislsvfslvmdrlvrfgrtravylasva 360

QY 361 AFPVAAGATCLSHSVAVVTASAAITGFTFSALQILPVTYTLASLYHREKOVFLPKYRGDTGG 420
DB 361 afpvaagatclshsvavvtasaa1tgftfsalqilpvttytlaslyhrekovflpkyr 420

QY 421 ASSSDSLMTSFLPGKPKCAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 asseds1mtsflpgkpkcapppnghvaggsgllppppalcgasacdvsvrvvvgeptea 480

QY 481 RVVPGRGICLDLAITLDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAAGLGLVAIFYATQ 540
DB 481 rvvpgrgicldla1ldsafllsqvapslfmg1siqvs1qsvtaymvsaaag1glvaifyatq 540

RESULT 2
AAW69385
ID AAW69385 standard; Protein; 553 AA.
XX
AC AAW69385;
XX
DT 08-DEC-1998 (first entry)
XX
DE Prostate tumour specific gene clone L1-12 protein.
XX
KW Prostate tumour specific gene; human; prostate cancer; detection;
KW therapy.
XX
OS Homo sapiens.
XX
PN W09837418-A2.
```

```
XX 27-AUG-1998.
PD
XX 25-FEB-1998; 98WO-US03690.
PF
XX 09-FEB-1998; 98US-0904809.
PR
XX 25-FEB-1997; 97US-0806596.
PR
XX 01-AUG-1997; 97US-0904809.
PR
XX (CORI-) CORIXA CORP.
PA
XX Dillion DC, Xu J;
PI
XX WPI; 1998-480805/41.
DR
XX N-PSDB; AAV58586.
DR
XX Novel human prostate specific tumour protein and fragments - useful
XX for detecting and treating prostate cancers
XX
XX Example 1; Page 87-89; 141pp; English.
XX
CC This sequence is encoded by a human prostate tumour specific gene, and
CC can be used in the method of the invention. The method is for detecting
CC prostate cancer comprising contacting a biological sample with an agent
CC able to bind an immunogenic portion of a prostate protein (such as
CC this protein sequence). An antibody which binds to an immunogenic
CC portion of the prostate protein, and the method can be used to detect,
CC monitor progression of, or treat prostate cancers. The antibody may
CC also be conjugated to a therapeutic agent for use in therapy of prostate
CC cancers.
XX
SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 19; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.2e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLRRKKAQOLLVNLITFGLEVCLAAGITVYPPLLLEVGVEEKFTMWLGIG 60
DB 1 mvqrlwvsrllrrhrkaqllvnlitfglevclaagityvpplllevgveekfntmvlgi 60

QY 61 PVLGVCVPLLGASDHWGRYGRRRPFIWALSIGLILSLFLIPRAGWLAGLPCDPRPL 120
DB 61 pvlglvcvplllgssadhwgrgyrrrrpfiwalsigllslflipragwlaglpcdprpl 120

QY 121 ELALLILGVGLDFCGOVCFPLEALLSDFRDPDHCROQAVSVYAFMISLGGCLGYLLPA 180
DB 121 elallilgvglldfcgvcftpleallsdfrdpdchrqaysvyafmisl9gclgyllpa 180

QY 121 elallilgvglldfcgvcftpleallsdfrdpdchrqaysvyafmisl9gclgyllpa 180
DB 121 elallilgvglldfcgvcftpleallsdfrdpdchrqaysvyafmisl9gclgyllpa 180

QY 181 IDWTSALAPYLGTOECFLGLTLTFLTCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240
DB 181 idwtsalapylgtqeecflglltfltcvaatl1lvaeeaa1gptepaeglsapslsph 240

QY 241 CCPCRLARFNLGALLPRHLQCCMRPTRLRLRFLVAELCSWALMTFTLYTDFVGEGL 300
DB 241 ccpcrlarfnlgallprhlqccmrptrlrrlrfvaelcswalmftftlytdfvgegl 300

QY 301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
DB 301 yqgvpraeptgearrhydegvrmgs1glflqcaislsvfslvmdrlvrfgrtravylasva 360

QY 361 AFPVAAGATCLSHSVAVVTASAAITGFTFSALQILPVTYTLASLYHREKOVFLPKYRGDTGG 420
DB 361 afpvaagatclshsvavvtasaa1tgftfsalqilpvttytlaslyhrekovflpkyr 420

QY 421 ASSSDSLMTSFLPGKPKCAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 asseds1mtsflpgkpkcapppnghvaggsgllppppalcgasacdvsvrvvvgeptea 480

QY 481 RVVPGRGICLDLAITLDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAAGLGLVAIFYATQ 540
DB 481 rvvpgrgicldla1ldsafllsqvapslfmg1siqvs1qsvtaymvsaaag1glvaifyatq 540
```

```
QY 541 VVFDKSLAKYSA 553
Db 541 vvfdksdlakysa 553

RESULT 3
AAB28527
ID AAB28527 standard; Protein; 553 AA.
AC AAB28527;
DT 07-FEB-2001 (first entry)
XX
DE Protein encoded by human breast tumour cDNA clone P501S.
XX
KW Human; breast tumour antigen; cytostatic; immunotherapy;
KW breast cancer; vaccine.
XX
OS Homo sapiens.
XX
PN WO200061756-A2.
XX
PD 19-OCT-2000.
XX
PF 10-APR-2000; 2000WO-US09688.
XX
PR 09-APR-1999; 99US-0288950.
XX
PR 02-JUL-1999; 99US-0346327.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Xu J, Dillon DC;
XX
XX WPI: 2000-638568/61.
DR N-PSDB; AAC79473.
XX
PT A novel isolated polypeptide comprising an immunogenic portion of a
PT breast cancer protein useful in the detection and treatment of breast
PT cancer -
XX
PS Claim 2; Page 92-93; 95pp; English.
XX
CC The present sequence is encoded by a cDNA sequence which was isolated
CC from a breast tumour cDNA library. It is provided in a specification
CC relating to compounds for immunotherapy and diagnosis of breast cancer.
CC Breast tumour antigens and the polynucleotides that encode them may be
CC used in the production of a pharmaceutical composition to be used in the
CC treatment of breast cancer. Proliferated T cells and incubated antigen
CC presenting cells are also required. The polypeptides and polynucleotides
CC may also be used to produce a vaccine.
XX
SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 21; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.2e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLVWSRLRRKQALLVNLTFGLVCLAAAGITYVPPDLLLEVGEVEKEFTMWVLGIG 60
Db 1 mvqrlvwsrlrrkqaillvnlftglvclaaagityvppdlllevgveekfntmvlgi 60
QY 61 PVGLVCPVLGASDHWGRYGRRRPFTWALSGLTLLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 pvglvcpvlgsadhwgrgyrrrrpftwalsgltsllslflipragwlagllcpdprpl 120
QY 121 ELALLILGVLDFCGQVCFPLEALLSDFRDPHCRQAYSVYAFMISLGGCLGYLPA 180
Db 121 elallilvgldfcgvcftpleallsdlfrdpdchrcqaysvyatfmslggcylgylpa 180
QY 181 IDWDTSAALAPYLGTOBECFLGLTLTFLTCVAATLLVAEEALGPTPEAGLSAPSLSPH 240
Db 181 idwdtasalapylgtobecflgltlftlftcvaatlvaeeaalgtptepaeglsapslsph 240

Db 181 idwtsalapylgtqecflgltlftlftcvaatlvaeeaalgtptepaeglsapslsph 240
QY 241 CCPCRARLAFRNLGALLPRLHOLCCRMPTLRRLFVAELCSWALMTFTLFTDFVGEGL 300
Db 241 ccpccrarlafrnlgallprlhqccrmptlrllfvaelcswalmntftlftdfvgegl 300
QY 301 YGVVPRABPGTEARRHYDEGVRMGLGLFLOCAISLVFSLVMDRLVQRTGTRAVYLASVA 360
Db 301 ygvvpraepgtearrhydegvrimgslglfqlcaislvsfslvmdrlvqrfgrtgravylasva 360
QY 361 AFPVAAGATCLSHSVAVVTASAAALTGFTFSALQIPLYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 afpvaagatclshsvavvtasaaaltgftfsalqilpytlaslyhrekqvlpkyrqdtgg 420
QY 421 ASSEDSLMTSFLPGPKPGAPFNGHVGAGSGLLPPPPALCGASACDVSRRVVGEPTGA 480
Db 421 assedslmtsflpgpkpgapfngvhgagsgllppppalcgasacdvsrrvvvgepte 480
QY 481 RVVPGRGICLDLAIDLSDAFLLSOVAPSLPMGSIIVOLSQSVTAYMVSAAAGLGLVAIFYATQ 540
Db 481 rvvpgrgicldlaidsafllsqvapslmgsvivqlsgsvtaymvsaaaglgivalfatq 540
QY 541 VVFDKSLAKYSA 553
Db 541 vvfdksdlakysa 553

RESULT 4
AAY82002
ID AAY82002 standard; Protein; 553 AA.
AC AAY82002;
DT 13-JUN-2000 (first entry)
XX
DE Human immunogenic prostate tumour protein sequence SEQ ID NO:113.
XX
KW Human; prostate cancer; diagnosis; tumour; gene therapy; detection;
KW immunogenic; cytostatic; vaccine.
XX
OS Homo sapiens.
XX
PN WO200004149-A2.
XX
PD 27-JAN-2000.
XX
PF 14-JUL-1999; 99WO-US15838.
XX
PR 14-JUL-1998; 98US-0115453.
PR 14-JUL-1998; 98US-0116134.
PR 23-SEP-1998; 98US-0159812.
PR 23-SEP-1998; 98US-0159822.
PR 15-JAN-1999; 99US-0232149.
PR 15-JAN-1999; 99US-0232880.
PR 09-APR-1999; 99US-0288946.
XX
PA (CORI-) CORIXA CORP.
XX
PI Dillon DC, Harlocker SL, Yuqiu J, Xu J, Mitcham JL;
XX
XX WPI: 2000-171268/15.
DR
PT New polypeptide useful for treating and diagnosing prostate cancer
PT comprises an immunogenic portion of prostate tumor protein -
XX
PS Claim 3; Page 138-139; 263pp; English.
XX
CC The present invention describes isolated polypeptides, comprising an
CC immunogenic portion of a prostate tumour protein (PTP). The polypeptides
CC and polynucleotides encoding them have cytostatic activity and can be
CC used in vaccines and in gene therapy. The polypeptides and
CC polynucleotides encoding them, antigen presenting cells which express
CC the polypeptides, antibodies against the polypeptides and vaccines
```

comprising them can be used for inhibiting the development of prostate cancer in a patient. The polypeptides can be used to generate antibodies or anti-idiotypic antibodies for passive immuno therapy. A portion of the polynucleotides encoding the polypeptides can be used as a probe or to modulate the expression of the polypeptides. AAA06241 to AAA06691 and AA82000 to AA82020 represent sequences used in the exemplification of the present invention.

XX Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 21; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.2e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQLWVSRLLRHRKAQLLVNLLTFGLVCLAAAGTYVPPLLLEVGVEEKFTMTVLGIG 60
Db 1 mvqlwvsrllrhrkaqlllvlltftglevclaagtyvpplllevgveekfntmvlgi 60
QY 61 PVGLVGVPLLGSDHWRGGRYRRPFIFWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 pvlglvcvpllgssdhwrgyrrrpfifwalsigillsflipragwlagllcpdprpl 120
QY 121 ELALLILGVGLDFCGOVCFPLEALLSDLFDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
Db 121 elallilgvglldfcgvcftpleallsdldfdrpdhcrqaysvyafmislggclgyllpa 180
QY 181 IDWTSALAPYLGTOECFLGLTLFLITCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
Db 181 idwtsalapylgtqecflgltlflitcvaatlvaeeaalgptepaeglsapslsph 240
QY 241 CCPCRLAFRNLCALLPRHLQCCMRPTRLRLFVAELCSWMAALMTFTLYTDFVGEGL 300
Db 241 ccpcrlarfnlrgallprhlqccmrptlrlfvaelcswmalmtftlytdfvgegl 300
QY 301 YQGVPRAEPTGTEARRHYDEGVRMGSGLGLFQCAISLVFSLVMDRLVORFQTRAVYLASVA 360
Db 301 yqgvpraepgtearrhydegvrmsgslglfqcalslvfslvmdrlvorfqtravylasva 360
QY 361 APFVAAGATCLSHSVAVVTASAALTGTFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 apfvaagatclshsvavvtasaaltgtftfsalqilpytlaslyhrekqvflpkyrtdtg 420
QY 421 ASSDLSMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
Db 421 assdlsmtsfpgpkpgapfngpnhvgaggsgllpppalcgasacdvsvrvvvgepta 480
QY 481 RVVPGRGICLDLAIDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIFYATQ 540
Db 481 rvpgrgicldlaidsafllsqvapslfmgsivqlsqsvtaymvsaaglglyvaifyatq 540
QY 541 VVFDKSDLAKYSA 553
Db 541 vvfdksdlakysa 553

RESULT 5

AAU04961
ID AAU04961 standard; Protein; 553 AA.

XX AAU04961;

XX AAU04961;

DT 24-OCT-2001 (first entry)

XX Human prostate tumour protein L1-12.

XX Human; prostate tumour protein; prostate cancer.

XX Homo sapiens.

XX US6262245-B1.

XX 17-JUL-2001.

XX 25-FEB-1998; 98US-0030607.
XX 25-FEB-1997; 97US-0806099.
PR 01-AUG-1997; 97US-0904804.
PR 09-FEB-1998; 98US-0020956.
XX (CORI-) CORIXA CORP.
XX Xu J, Dillon DC;
XX WPI: 2001-440862/47.
DR N-PSDB; AAS10108.
XX Novel polynucleotide encoding polypeptide comprising a portion of
PT prostate tumour protein useful for inhibiting development of prostate
PT cancer or for treating prostate cancer in a patient.
XX Example 1; Column 125-127; 105pp; English.
XX The sequence is a partial prostate tumour protein, encoded by a prostate
CC tumour specific cDNA. The DNA is useful for inhibiting the development
CC of prostate cancer or for treating prostate cancer in a patient.
XX Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 22; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.2e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQLWVSRLLRHRKAQLLVNLLTFGLVCLAAAGTYVPPLLLEVGVEEKFTMTVLGIG 60
Db 1 mvqlwvsrllrhrkaqlllvlltftglevclaagtyvpplllevgveekfntmvlgi 60
QY 61 PVGLVGVPLLGSDHWRGGRYRRPFIFWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 pvlglvcvpllgssdhwrgyrrrpfifwalsigillsflipragwlagllcpdprpl 120
QY 121 ELALLILGVGLDFCGOVCFPLEALLSDLFDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
Db 121 elallilgvglldfcgvcftpleallsdldfdrpdhcrqaysvyafmislggclgyllpa 180
QY 181 IDWTSALAPYLGTOECFLGLTLFLITCVAATLLVAEEAALGPTPEAGLSAPLSLSPH 240
Db 181 idwtsalapylgtqecflgltlflitcvaatlvaeeaalgptepaeglsapslsph 240
QY 241 CCPCRLAFRNLCALLPRHLQCCMRPTRLRLFVAELCSWMAALMTFTLYTDFVGEGL 300
Db 241 ccpcrlarfnlrgallprhlqccmrptlrlfvaelcswmalmtftlytdfvgegl 300
QY 301 YQGVPRAEPTGTEARRHYDEGVRMGSGLGLFQCAISLVFSLVMDRLVORFQTRAVYLASVA 360
Db 301 yqgvpraepgtearrhydegvrmsgslglfqcalslvfslvmdrlvorfqtravylasva 360
QY 361 APFVAAGATCLSHSVAVVTASAALTGTFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 apfvaagatclshsvavvtasaaltgtftfsalqilpytlaslyhrekqvflpkyrtdtg 420
QY 421 ASSDLSMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
Db 421 assdlsmtsfpgpkpgapfngpnhvgaggsgllpppalcgasacdvsvrvvvgepta 480
QY 481 RVVPGRGICLDLAIDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIFYATQ 540
Db 481 rvpgrgicldlaidsafllsqvapslfmgsivqlsqsvtaymvsaaglglyvaifyatq 540
QY 541 VVFDKSDLAKYSA 553
Db 541 vvfdksdlakysa 553

RESULT 6

QY	ID	AAAM01117 standard; Protein; 553 AA.	DE	QY	ID	CCPCRARLAFNRNLGALLPRLHOLCCMRPTLRLRLFAELCSNMALMTFTLFTDFVGEGL	300
XX	AAAM01117		XX	241	CCPCRARLAFNRNLGALLPRLHOLCCMRPTLRLRLFAELCSNMALMTFTLFTDFVGEGL	300	
XX	AAAM01117		Db	241	CCPCRARLAFNRNLGALLPRLHOLCCMRPTLRLRLFAELCSNMALMTFTLFTDFVGEGL	300	
XX	04-OCT-2001 (first entry)		QY	301	YGVPRADPGTEARRHYDEGVRMGSLGLFLQCAISLVSFLVMDRLVQRFQTRAVYLASVA	360	
XX	Human prostate-specific amino acid sequence L1-12.		Db	301	YGVPRADPGTEARRHYDEGVRMGSLGLFLQCAISLVSFLVMDRLVQRFQTRAVYLASVA	360	
XX	Human; prostate cancer; prostate-specific; diagnosis; vaccine;		QY	361	AFPVAAGATCISHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG	420	
XX	cytostatic; gene therapy; metastasis.		Db	361	AFPVAAGATCISHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG	420	
OS	Homo sapiens.		QY	421	ASSEDLSMTSLPGLPKGAPFPNGHVAGSGLLPPPALCCASACDVSRRVVVGEPTEA	480	
XX	WO200151633-A2.		Db	421	ASSEDLSMTSLPGLPKGAPFPNGHVAGSGLLPPPALCCASACDVSRRVVVGEPTEA	480	
XX	19-JUL-2001.		QY	481	RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAAGLGLVATYFATQ	540	
XX	16-JAN-2001; 2001WO-US01574.		Db	481	RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAAGLGLVATYFATQ	540	
XX	14-JAN-2000; 2000US-0483672.		QY	541	VVFDKSDLAKYSA 553		
XX	(CORI-) CORIXA CORP.		Db	541	VVFDKSDLAKYSA 553		
XX	Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;		RESULT	7			
XX	Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;		AAAG99002				
XX	Wang A, Meagher MJ;		ID	AAAG99002 standard; Protein; 553 AA.			
XX	WPI; 2001-425873/45.		XX	AC	AAAG99002;		
XX	New polynucleotide encoding a prostate-specific protein, for		XX	DT	25-SEP-2001 (first entry)		
XX	diagnosing, monitoring and treating prostate cancer in a patient and		XX	DE	Human prostate-specific amino acid sequence L1-12/P501S.		
XX	for use in vaccines.		XX	XX	Human; prostate cancer; therapy; diagnosis; cat eye syndrome;		
XX	Claim 2; Page 267-268; 543pp; English.		XX	XX	chromosome 22q11.2; prostate-specific protein; Chromosome 1;		
XX	The present invention describes polynucleotide sequences (I) which encode		XX	XX	prostate specific antigen; PSA.		
XX	prostate-specific proteins (II). (I) and (II) have cytostatic activity,		XX	OS	Homo sapiens.		
XX	and can be used in vaccine production and gene therapy. (I), (II),		XX	PN	WO2001134802-A2.		
XX	antibodies to (II), fusion proteins comprising (II), and isolated		XX	PD	17-MAY-2001.		
XX	T cells prepared using (I) or (II) are used to treat cancer in a patient.		XX	PF	09-NOV-2000; 2000WO-US30904.		
XX	(I) and the antibodies are also used in the detection of cancer in a		XX	PR	12-NOV-1999; 99US-0439313.		
XX	patient. The cancer that is diagnosed or treated is particularly		XX	PR	18-NOV-1999; 99US-0443686.		
XX	prostate cancer. (I) and (II) can be used in vaccines. The antibodies or		XX	PA	(CORI-) CORIXA CORP.		
XX	(I) can be used for monitoring the progression of cancer in a patient.		XX	PI	Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;		
XX	(I) and (II) can also be used to improve diagnostic and therapeutic		XX	PI	Kalos MD, Retter MW, Stolk JA, Day CH, Skeiky YAW, Wang A;		
XX	methods for prostate cancer. They can indicate the level of metastasis		XX	DR	WPI; 2001-308785/32.		
XX	as well as the prostate volume. AA93357 to AA93944 and AA01115 to		XX	XX	Isolated polypeptide comprising at least an immunogenic portion of a		
XX	AA01318 represent polynucleotide and amino acid sequences used in the		XX	PT	prostate-specific protein, useful in the diagnosis and therapy of		
XX	exemplification of the present invention.		XX	PT	prostate cancer.		
XX	Sequence 553 AA;		XX	PS	Claim 3; Page 167-168; 325pp; English.		
XX	Query Match 100.0%; Score 2861; DB 22; Length 553;		XX	XX	The present invention describes an isolated polypeptide (P1) comprising		
XX	Best Local Similarity 100.0%; Pred. No. 1.2e-269;		XX	CC	at least an immunogenic portion of a prostate-specific protein, or its		
XX	Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		XX	CC	variant. Also described are polynucleotides (N1) encoding (P1). (P1)		
QY	1 MVQRLWSRLRHRKAQLLVNLTFTGLEVLAAGITVPPPLLELVGVESEKFTMWLGIG 60		XX	CC	(N1) have cytostatic activity and can be used in vaccine production.		
Db	1 mvqrlwsrllrhrkaqlvllnltftglevclaagityvppllilevgveekfntmvlgi 60		XX	CC	The polypeptides, nucleic acids and antibodies from the present		
QY	61 PVLGIVCPVLGASDHWGRGRRPFIWALSIGLILLSFLIPRAGWLAGLILCPDRPL 120		XX	CC	invention are useful in the diagnosis and therapy of prostate cancer.		
Db	61 pvlgivcpvlgsasdhwrgrrpfiwalslgillsflipragwlaglilcpdrpl 120		XX	CC	Prostate specific genes P704P, P712P, P747P, P753P and B305D are located		
QY	121 ELALLILGVLLDFCGVCFTEALLSDLFDPDHCRCQAYSVYAFMISLGGCLGVLLPA 180		XX	CC	in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome		
Db	121 elallilgvlldfcgvcftleallsdldfdrpdhrcqaysvafmislvgclglp 180		XX	CC	region. Prostate specific antigen (PSA) 501S was located on		
QY	181 IDWTSALAPYLTQECLEGLITLFTLCTCAATLLVAEEAALGPTEPAEGLSAPLSPH 240						
Db	181 idwtsalapyltqgeclfglltftlctvaatlivaeeaalq						

```
CC chromosome 1, AAH84671 to AAH85143 and AAC99000 to AAC99077 represent
CC polynucleotide and polypeptide sequences used in the exemplification
CC of the present invention.
XX
SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 22; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.2e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRWVSRLLRRHKAQLLVNLTFTGLEVCIAAGIYVPPPLLEVGVEEKMTWLGIG 60
DB 1 mvqrwvsrllrrhkaqlllvnlftglevciaagityvpplllevgveekfntmvlgi 60

QY 61 PVGLVCVPLLSASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPPRPL 120
DB 61 pvglvcvplllsgasdhwgrgyrrrrpfiwalslgillsflipragwlagllcpdprl 120

QY 121 ELALLIIGVGLDFCGOVCFPTLEALLSDFRPDHCROAYSVYAFMISLGGCLGYLLPA 180
DB 121 elallliigvlldfcgvvcftpleallsdfrpdhcrqaysvyafmislggclgyllpa 180

QY 181 IDWTSALAPYLGTOECFLGLTLTFLTCVAATLLVAEEAALGPTEPAEGLSAPLSPH 240
DB 181 idwtsalapylgtqecflgltlftlftcvaatllvaeaaalgppeaeglsapslsph 240

QY 241 CCPCRARLAFNLGALLPRLHQLCCRPTRLRLFVAELCSWMLMTFTFLFYDFVGEGL 300
DB 241 ccpcrarlafnlgallprlhqlccrptlrrlrfvaelcswmlmtftflfydfvgegl 300

QY 301 YQGVPRAEPTGEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLSVA 360
DB 301 yqgvpraepgtearrhydegvrmsgslflqcaislvsfslvmdrlvqrfctravylsava 360

QY 361 AFPVAGATCLSHVAVVTASAALTGFTFSAQLIPLYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 afpvaagatclshsavvtasaaltgftfsaqlilpytlaslyhrekqvfllpkrygdtgg 420

QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 asseedsimtsflpgpkpgapfngnhvagsgllppppalcgasacdvsvrvvvgepte 480

QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVWSAAGLGLVATYFATQ 540
DB 481 rvpvprgicldlaildsafllsqvapslfmgslvqlsqsvtaymvwsaaglglvatyfatq 540

QY 541 VVFDKSDLAKYSA 553
DB 541 vvfdksdlakysa 553

RESULT 8
AAG62150
ID AAG62150 standard; Protein; 553 AA.
XX
AC AAG62150;
XX
DT 06-JUL-2001 (first entry)
XX
DE Human P501S inventive antigen SEQ ID NO: 333.
XX
KW Human; mouse; immunotherapy; cancer; leukaemia; Wt1; Wilm's tumour gene;
KW chromosome 11p13; zinc finger transcription factor.
XX
OS Homo sapiens.
XX
PN WO200125273-A2.
XX
PD 12-APR-2001.
XX
PF 04-OCT-2000; 2000WO-US27465.
XX
```

```
PR 04-OCT-1999; 99US-0157459.
XX (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Xu J, Cheever MA, Reed SG;
XX
DR WPI; 2001-328324/34.
XX
PT Polypeptide comprising part of the Wilm's Tumour gene product sequence is
PT used in the diagnosis and treatment of malignant diseases e.g. leukemia
PT and cancer associated with Wt1
XX
PS Disclosure; Page 212-213; 228pp; English.
XX
CC The present invention describes compositions comprising peptides derived
CC from the Wilm's tumour protein Wt1 and methods for their use in treating
CC malignant diseases. Peptides derived from both the murine and human Wt1
CC proteins are provided. The human Wt1 gene is found on chromosome 11p13,
CC and the protein was shown to be a zinc finger transcription factor. The
CC immunogenic peptides of the invention are particularly useful in the
CC diagnosis and treatment of cancer and leukaemia. The present sequence is
CC a polypeptide described in the exemplification of the invention.
XX
SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 22; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.2e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRWVSRLLRRHKAQLLVNLTFTGLEVCIAAGIYVPPPLLEVGVEEKMTWLGIG 60
DB 1 mvqrwvsrllrrhkaqlllvnlftglevciaagityvpplllevgveekfntmvlgi 60

QY 61 PVGLVCVPLLSASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPPRPL 120
DB 61 pvglvcvplllsgasdhwgrgyrrrrpfiwalslgillsflipragwlagllcpdprl 120

QY 121 ELALLIIGVGLDFCGOVCFPTLEALLSDFRPDHCROAYSVYAFMISLGGCLGYLLPA 180
DB 121 elallliigvlldfcgvvcftpleallsdfrpdhcrqaysvyafmislggclgyllpa 180

QY 181 IDWTSALAPYLGTOECFLGLTLTFLTCVAATLLVAEEAALGPTEPAEGLSAPLSPH 240
DB 181 idwtsalapylgtqecflgltlftlftcvaatllvaeaaalgppeaeglsapslsph 240

QY 241 CCPCRARLAFNLGALLPRLHQLCCRPTRLRLFVAELCSWMLMTFTFLFYDFVGEGL 300
DB 241 ccpcrarlafnlgallprlhqlccrptlrrlrfvaelcswmlmtftflfydfvgegl 300

QY 301 YQGVPRAEPTGEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVQRFCTRAVYLSVA 360
DB 301 yqgvpraepgtearrhydegvrmsgslflqcaislvsfslvmdrlvqrfctravylsava 360

QY 361 AFPVAGATCLSHVAVVTASAALTGFTFSAQLIPLYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 afpvaagatclshsavvtasaaltgftfsaqlilpytlaslyhrekqvfllpkrygdtgg 420

QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 asseedsimtsflpgpkpgapfngnhvagsgllppppalcgasacdvsvrvvvgepte 480

QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVWSAAGLGLVATYFATQ 540
DB 481 rvpvprgicldlaildsafllsqvapslfmgslvqlsqsvtaymvwsaaglglvatyfatq 540

QY 541 VVFDKSDLAKYSA 553
DB 541 vvfdksdlakysa 553

RESULT 9
AAB74800
```

ID	AAB74800 standard; Protein; 553 AA.		
XX	AAB74800;		
AC			
XX			
DT	14-JUN-2001 (first entry)		
XX			
DE	Prostate tumour antigen predicted amino acid sequence for L1-12.		
XX			
KW	Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;		
KW	prostate cancer; immunogenic; cytostatic; vaccine.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200125272-A2.		
XX			
PD	12-APR-2001.		
XX			
PF	04-OCT-2000; 2000WO-US27464.		
XX			
PR	04-OCT-1999; 99US-0157455.		
XX			
PA	(CORI-) CORIXA CORP.		
XX			
PI	Xu J, Skeiky YAW, Reed SG, Cheever MA;		
XX			
DR	WPI; 2001-245062/25.		
DR	N-PSDB; AAH02530.		
XX			
PT	Prostate specific protein and its encoding polynucleotide, useful for		
PT	the treatment and diagnosis of prostate cancer -		
XX			
PS	Claim 3; Page 157-158; 276pp; English.		
XX			
CC	The present invention describes an isolated polypeptide (I) comprising		
CC	at least an immunogenic portion of a prostate tumour antigen protein or		
CC	its variant. (I) have cytostatic activity and can be used in vaccine		
CC	production. (II), prostate tumour antigen polynucleotides, an antigen		
CC	presenting cell (APC e.g. a dendritic cell) that expresses (I), and a		
CC	pharmaceutical composition containing (I) are useful for inhibiting the		
CC	development of cancer in a patient. Antibodies specific for prostate		
CC	specific proteins and oligonucleotides that hybridise to a		
CC	polynucleotide that encodes a prostate specific protein are useful		
CC	for detecting the presence or absence of a cancer or monitoring the		
CC	progression the progression of a cancer, especially prostate cancer.		
CC	AAH02422 to AAH2872, AAB74798 to AAB74821 and AAB74830 are sequences		
CC	used in the exemplification of the present invention.		
XX			
SQ	Sequence 553 AA;		
Query Match 100.0%; Score 2861; DB 22; Length 553;			
Best Local Similarity 100.0%; Pred. No. 1.2e-269;			
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	MVQRLLWSRLRHRKAQLLVNLTTCLEVCAGITVPPPLLEVGVEKEKFTMTVLGIG	60
Db	1	mvqrllwsrlrhrkaqlllvnlttfglevclagagityvppllllevgveekftmtvlgig	60
QY	61	PVLGLVCVPLLGASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDRPL	120
Db	61	pvlglvcvpllgasdhwgrgyrrrrpfiwalslgillsllflpragwlagllcpdrpl	120
QY	121	ELALLIILGVLLDFCGVCTPTEALLSLDFRDPDHCROAYSVAWMSISGGCLGYLLPA	180
Db	121	elallilgvlldfcgvctpteallslldfrdpdchrqaysvaawmsisggclgyllpa	180
QY	181	IDWDTSAALPYLGTQECLEGLTLIFLTCAATLLVAEEAALGPTPEAGLSAPSLSPH	240
Db	181	idwdsalapylgtqeclglltliftcvaatlvaeeaalgptepaeglsapslsph	240
QY	241	CCPCRARLAFRNIGALLPRHQCCMRPTRLRLFVAELCSWMLMFTFLFYTFDFVGEGL	300
Db	241	ccpcrarlafnrigallprhqccmrptrlrflvaelcswmlmftflfytdfvgegl	300
QY	301	YQGVPRAPGTEARRHYDEGVRMGSLGLEFLQCAISLVSFLVMDRLVQREGTRAVYLASVA	360
Db	301	yqgvpraepgtearrhydegvmgslglflqcaislvslvmdrlvqrgtravylasva	360
QY	361	AFPVAAGATCCLSHSVAVVTASAAALGFTFSALQILPYTTLASLYHREKQVFLPKYRGDTGG	420
Db	361	afpvaagatclshsvavvtasaaaltgftfsalqilpyttlaslyhrekqvlfpkyrgdtgg	420
QY	421	ASSEDSLMTSFLPGPKPGAPFPNGHVAGSGLLPPPPALCGASACDVSVRVVVGEPTEA	480
Db	421	assedslmtsflpgpkpgapfngvhvagsgllppppalcgasacdvsvrvvvgepte	480
QY	481	RVVPRGICLDLAIDLDSAFELLSQVAPSLFMGSIVQLSQSVTAYVMVSAAGLGLVAIFPATQ	540
Db	481	rvvpgrgicldlaidsafllsqvapslfmgsivqlsqsvtaymvsaaglglvaiyfatq	540
QY	541	VVFDRKSLAKYSA 553	
Db	541	vvdksdlakysa 553	
RESULT 10			
AAB74830			
ID	AAB74830 standard; Protein; 1079 AA.		
XX			
AC	AAB74830;		
XX			
DT	14-JUN-2001 (first entry)		
XX			
DE	Prostate tumour antigen amino acid sequence for a fusion protein.		
XX			
KW	Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;		
XX	prostate cancer; immunogenic; cytostatic; vaccine.		
OS	Homo sapiens..		
XX			
PN	WO200125272-A2.		
XX			
PD	12-APR-2001.		
XX			
PF	04-OCT-2000; 2000WO-US27464.		
XX			
PR	04-OCT-1999; 99US-0157455.		
XX			
PA	(CORI-) CORIXA CORP.		
XX			
PI	Xu J, Skeiky YAW, Reed SG, Cheever MA;		
XX			
DR	WPI; 2001-245062/25.		
XX			
PT	Prostate specific protein and its encoding polynucleotide, useful for		
PT	the treatment and diagnosis of prostate cancer -		
XX			
PS	Disclosure; Page 272-276; 276pp; English.		
XX			
CC	The present invention describes an isolated polypeptide (I) comprising		
CC	at least an immunogenic portion of a prostate tumour antigen protein or		
CC	its variant. (I) have cytostatic activity and can be used in vaccine		
CC	production. (II), prostate tumour antigen polynucleotides, an antigen		
CC	presenting cell (APC e.g. a dendritic cell) that expresses (I), and a		
CC	pharmaceutical composition containing (I) are useful for inhibiting the		
CC	development of cancer in a patient. Antibodies specific for prostate		
CC	specific proteins and oligonucleotides that hybridise to a		
CC	polynucleotide that encodes a prostate specific protein are useful		
CC	for detecting the presence or absence of a cancer or monitoring the		
CC	progression the progression of a cancer, especially prostate cancer.		
CC	AAH02422 to AAH2872, AAB74798 to AAB74821 and AAB74830 are sequences		
CC	used in the exemplification of the present invention.		
XX			
SQ	Sequence 1079 AA;		

Query Match 100.0%; Score 2861; DB 22; Length 1079;
Best Local Similarity 100.0%; Pred. No. 3e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVORLWVSRLRRRAQLLVNLTFLGLEVCIAAGTYVPPDLLLEVGVEEKFTMTVLGIG 60
DB 527 mvqrlwvsrlrrhkaqlilvnlftflglevciaagityvppdlllevgvveekfntmvlgi9 586
QY 61 PVLGLVCVPLLSASDHWGRYGRPPFTWALSGLILLSLFLIPRAGWLAGLLCPDPRPL 120
DB 587 pvlglvcvpllsasdhwgrgyrrprfwiwalsgillsflipragwlagllcpdprpl 646
QY 121 ELALLILVGLIDFCQGVCFPLEALLSDFRDPDRCROAYSVYAFMISLGGCLGILLPA 180
DB 647 elallilvgliidfcgvqcfpleallsdfrdpdrcrdaysvyafmislggcilgylpa 706
QY 181 IDWDTLSALAPYLGTOECLFGLITLTLFLTCVAATLLVABEALGPTPEAGLSAPLSLPH 240
DB 707 idwdsalapylgtqeclfgliltlfltcvaatlivaeeaalgtpeaeglsapslsph 766
QY 241 CQPCRLAFRNLGALLPRHLQCCRMPTLRLFLVAELCSWMLMTFTFLFYTDFVGEGL 300
DB 767 cpcrclafalnlgallprhlqccrmptlrlfvaelcswmlmtftflfytdfvegl 826
QY 301 YQGVPRAEPTGTEARRHYDEGVRMGSGLGLFQCAISLVSFLVMDRLVQRFCTRAVYLASVA 360
DB 827 yqgvpraepgtearrhydegvrmgsglglfqcailsvflvmdrlvqrfctravylasva 886
QY 361 AFPVAAGATCLSHSNAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 887 afpvaagatclshsvavvtasaaltgtftsalsqilpytlaslyhrekqvlpkrygdtgg 946
QY 421 ASSEDSLMTSFLPGPKPGAPFNGHVGAGSGLLPPLPPALCGASACDVSRRVVVGTEA 480
DB 947 asseedsimtsflpgpkpgapfngvhvgagsgllppppalcgasacdvsrrvvvgteea 1006
QY 481 RVVPGRGICLDLAIDLSAFLLSOVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIYFATQ 540
DB 1007 rvpvgrgicldlaidsaflsqvapslfmgsivqlsqsvtaymvsaaglglvaiyfataq 1066
QY 541 VVFDKSDLAKYSA 553
DB 1067 vvfkdslakysa 1079

RESULT 11
ID AAM01318 standard; Protein; 595 AA.
AC AAM01318;
XX AAM01318;
DT 04-OCT-2001 (first entry)
XX Alpha prepro-p501S recombinant protein amino acid sequence.
XX Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW cytostatic; gene therapy; metastasis.
XX Homo sapiens.
XX WO200151633-A2.
XX 19-JUL-2001.
XX 16-JAN-2001; 2001WO-US01574.
XX 14-JAN-2000; 2000US-0483672.
XX (CORI-) CORIXA CORP.
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;
PI Wang A, Meagher NJ;

XX WPI; 2001-425873/45.
XX New polynucleotide encoding a prostate-specific protein, for
PT diagnosing, monitoring and treating prostate cancer in a patient and
PT for use in vaccines -
XX Example 17; Page 541-543; 543pp; English.
XX The present invention describes polynucleotide sequences (I) which encode
CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC and can be used in vaccine production and gene therapy. (I), (II),
CC antibodies to (II), fusion proteins comprising (II), and isolated
CC T cells prepared using (I) or (II) are used treat cancer in a patient.
CC (I) and the antibodies are also used in the detection of cancer in a
CC patient. The cancer that is diagnosed or treated is particularly
CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC (I) can be used for monitoring the progression of cancer in a patient.
CC (I) and (II) can also be used to improve diagnostic and therapeutic
CC methods for prostate cancer. They can indicate the level of metastasis
CC as well as the prostate volume. AA93357 to AA93944 and AAM01115 to
CC AAM01318 represent polynucleotide and amino acid sequences used in the
CC exemplification of the present invention.
XX Sequence 595 AA:

Query Match 90.7%; Score 2596; DB 22; Length 595;
Best Local Similarity 95.1%; Pred. No. 8.1e-244;
Matches 501; Conservative 9; Mismatches 9; Indels 8; Gaps 1;

QY 35 GIVVPPPLLVEGVVEKEFMT-----MVLGIGPVILGVCVPLLSASDHWGRYGRRR 86
DB 61 glifinttiasiaakeegvslekreaeamvlgigpvlgvcvpllsasdhwgrgyrrr 120
QY 87 PFTWALSGLILLSLFLIPRAGWLAGLLCPDPRPLEALLILVGLIDFCQGVCFPLEAL 146
DB 121 pfiiwalsglillsflipragwlagllcpdprpleallilvgliidfcgvqcfpleal 180
QY 147 LSDLFRDPDRCROAYSVYAFMISLGGCLGYLLPAIDWDTLSALAPYLGTOECLFGLITLI 206
DB 181 lsdlfrdpdrcrdaysvyafmislggclyllpaidwdsalapylgtqeclfgliltli 240
QY 207 FLTCVAATLLVABEALGPTPEAGLSAPLSPHCCPCRLAFRNLGALLPRHLQCCCR 266
DB 241 fltcvaatlivaeeaalgtpeaeglsapslshccpcrclafalnlgallprhlqccr 300
QY 267 MPRTLRLFLVAELCSWMLMTFTFLFYTDFVGEGLYQGVPRAEPTGTEARRHYDEGVRMGS 326
DB 301 mprtlrlfvaelcswmlmtftflfytdfvgelyqgvpraepgtearrhydegvrmgsl 360
QY 327 GLFLQCAISLVSFLVMDRLVQRFCTRAVYLASVAAPFVAAGATCLSHSNAVVTASAAITG 386
DB 361 glflqcaislvflvmdrlvqrfctravylasvaafpvaagatclshsvavvtasaaltg 420
QY 387 FTFSALQILPYTLASLYHREKQVFLPKYRGDTGGASEDSLMTSFLPGPKPGAPFNGHV 446
DB 421 ftfsalqilpytlaslyhrekqvlpkryrgdtggasedslmtsflpgpkpgapfnghv 480
QY 447 GAGSGLLPPLPPALCGASACDVSRRVVVGPTTEARVYVPGRGICLDLAIDLSAFLLSQVAP 506
DB 481 gagsgllppppalcgasacdvsrrvvvgpttearvvpgrgicldlaidsaflsqvap 540
QY 507 SLFMGSIVQLSQSVTAYMVSAAGLGLVAIYFATQVFDKSDLAKYSA 553
DB 541 slfmgsivqlsqsvtaymvsaaglglvaiyfataqvfdkdsdlakysa 587

RESULT 12
ID AAM01230
XX AAM01230 standard; Protein; 371 AA.
XX AAM01230;
AC AAM01230;

CC	angiogenic disorders, kidney disorders, gastrointestinal disorders,
CC	pregnancy-related disorders, endocrine disorders, and infections. The
CC	proteins can also be used to aid wound healing and epithelial cell
CC	proliferation, to prevent skin aging due to sunburn, to maintain organs
CC	before transplantation, for supporting cell culture of primary tissues,
CC	to regenerate tissues, to identify their cognate ligands or binding
CC	partners, and in chemotaxis, and can be used as a food additive or
CC	preservative to modify storage properties. Antibodies specific for a
CC	protein of the invention can be used in alleviating symptoms associated
CC	with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC	radioimmunoassay or enzyme linked immunosorbent assay (ELISA).
CC	The present sequence represents a human secreted protein of
CC	the invention.
XX	
XX	Sequence 371 AA;
SQ	
	Query Match 49.5%; Score 1416; DB 22; Length 371;
	Best Local Similarity 98.5%; Pred. No. 2.8e-129;
	Matches 263; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY	54 TMVLGIGPVGLVCVPLILGSASDHWGRGYGRRRPTFNALSGLILLSFLIPRAGWLAGLL 113 :: :
Db	28 silagigpvlgvcvplilgsasdhwgrgyrrrpfwiwalsgilillsflipragwlagll 87 :: :
QY	114 CPDRPLELALLIIGVGILLDFCGOVCTFPLEALLSDLFRDPDHCRQAYSVYAFMISLGGC 173
Db	88 cpdrpletalligvgilldfcgvctfpleallsdlfrdpdchrqaysvyafmisi9gc 147
QY	174 LGYLLPAIDNPTSALAPYLGTQECLFGLLTFLITCVAATLLVAEEAALGTFEPAEGLS 233
Db	148 lgyllpaidwtasalapylgtqeclfglltflitcvaatllvaeaaalgtepaegls 207
QY	234 APSLSPHCCPCRARLPENLCALLPRHLQCCRMPRTLRRLFEVALCSWMALMTFTLEYT 293
Db	208 apslsphccpcrarlafnlgallprlhqlccmrptlririfvaelcswmalmfttlift 267
QY	294 DFVGEGLYGVPRAEPGTERRHYDEG 320
Db	268 dfvgeglygvpraepgtearrhydeg 294
RESULT 14	
AAM01262	ID AAM01262 standard; Protein; 400 AA.
XX	AAM01262;
XX	04-OCT-2001 (first entry)
DT	Rall2-P501S-E2 construct amino acid sequence.
DE	Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW	cytostatic; gene therapy; metastasis.
KW	Homo sapiens.
OS	WO200151633-A2.
PN	19-JUL-2001.
XX	16-JAN-2001; 2001WO-US01574.
PF	14-JAN-2000; 2000US-0483672.
XX	(CORI-) CORIXA CORP.
XX	Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI	Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;
PI	Wang A, Meagher MU;
XX	WPI; 2001-425873/45.
DR	
XX	

PA (ABBO) ABBOTT LAB.
XX
PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;
PI Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;
PI Roberts-Rapp L, Russell JC, Stroupe SD;
XX
DR WPI; 1999-034731/03.
DR N-PSDB; AAV71181.
XX
PT New isolated prostate-specific polynucleotides - used to develop
PT products for the diagnosis and treatment of prostate diseases, e.g.
PT benign hyperplasia, prostatic or prostate cancer
XX
PS Claim 17; Pages 99-100; 122pp; English.
XX
CC The present sequence is encoded by the consensus sequence for a PS108
CC gene. The sequence is derived from overlapping clones AAV71166-79. The
CC clone sequences are PS108 gene-specific. They are used in the method
CC of the invention. The specification describes a method for detecting the
CC presence of a target PS108 polynucleotide in a test sample. The method
CC comprises contacting the test sample with at least 1 PS108-specific
CC polynucleotide or complement, and detecting the presence of the target
CC PS108 polynucleotide. The products can be used for detecting,
CC diagnosing, staging, monitoring, prognosticating, in vivo imaging, or
CC preventing or treating, or determining predisposition to diseases or
CC conditions of the prostate such as benign prostatic hyperplasia (BPH),
CC prostatitis, prostatic intraepithelial neoplasia (PIN) and cancer. In
CC particular the products can be used in drug screening and gene therapy.
XX
SQ Sequence 255 AA;

Query Match 45.0%; Score 1287; DB 20; Length 255;
Best Local Similarity 100.0%; Pred. No. 5.9e-117;
Matches 255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 GLYGVPRAPGTEARRHYDEGVNMGSLGLFLQCAISLVFSLVMDRLVQRFGRVAVLAS 358
Db 1 glygvprapgtearrhydegvmgsigiflqcaislvsfslvmdrlvqrfgrtravylas 60

QY 359 VAAPFVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDT 418
Db 61 vaafpvaagatclshsvavvtasaaltgftfsalqilpytlaslyhrekqvflpkyrgrdt 120

QY 419 GGASEDSLMTSFLPGPKGAPPNGHVAGGGLPPPPALCGASACDYSRVVVVGEP 478
Db 121 ggasedslmtsflpgpkgapngnhv9ag9sglppppalcgasacdsvrvvvvgsept 180

QY 479 EARVVPGRGICLDLAIDSAFLLSQVAPSLFMGSIVQLSQSVTAYMYSAAGLGLVAYFA 538
Db 181 earvvpgrgicldlaidsaflsqvapslfmgsivqlsqsvtaymvsaaaglgivayfa 240

QY 539 TQVFDKSDLAKYSA 553
Db 241 tqvfdksdlakysa 255

Search completed: February 26, 2002, 07:56:35
Job time: 225 sec